COMBINING NEXT-GENERATION PHENOTYPING AND GENOME-WIDE ASSOCIATION ANALYSIS TO EXPLORE THE GENETIC ARCHITECTURE OF NUTRIENT ACQUISITION IN DOMESTICATED ASIAN RICE (*Oryza Sativa* L.).

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Joshua Nathaniel Cobb

May 2013

COMBINING NEXT-GENERATION PHENOTYPING AND GENOME-WIDE ASSOCIATION ANALYSIS TO EXPLORE THE GENETIC ARCHITECTURE OF NUTRIENT ACQUISITION IN DOMESTICATED ASIAN RICE (*Oryza sativa* L.).

Joshua Nathaniel Cobb, Ph.D.

Cornell University 2013

The acquisition of inorganic matter from soil is fundamental to the survival of all plant life on earth, and is of particular concern regarding agriculturally relevant species. The summation of the inorganic constituency of an organism is collectively known as the ionome. The study of the ionome, which is called ionomics, has the potential to impact major agricultural concerns facing the planet such as nutrient use efficiency, heavy metal toxicity, soil salinification, bio-fortificiation, and exclusion of toxic metals or metalloids from the human diet. Partly due to the challenge of phenotyping such a complex network of correlated phenotypes amid tremendous environmental variation, the genetic architecture of the rice ionome is poorly characterized. This dissertation explores the genetic underpinnings of the phenotypic variation of the rice ionome and demonstrates the tractability of using controlled environments to survey large cross-sections of genetic variation. Genetic analysis was performed using both linkage and association mapping on a diverse and well-characterized bi-parental recombinant inbred line mapping population and panel of 400 diverse *Oryza sativa* accessions collected from around the world, respectively. The analysis identified 362 significant regions of the rice genome controlling the accumulation of 24 mineral elements in rice roots and shoots under hydroponic growth conditions. Seventy nine of these significant regions co-localize with known and putative candidate genes. Haplotypes

at candidate genes for shoot molybdenum and shoot sodium concentration along with physiological and gene expression experimentation validates the hypothesis that these candidates are likely to be responsible for a large portion of the phenotypic variance present in our populations. Furthermore we demonstrate that the *aus* sub-population excludes significantly more sodium from the shoot tissue and accumulates significantly more molybdenum than any other sub-population, and especially *indica*. Admixture analysis identified that *aus* haplotypes at candidate genes identified by the genetic mapping results significantly increase molybdenum shoot concentration and decrease sodium shoot concentration when present in otherwise *indica*-like genomes. The results from this study form the basis for extensive follow-up research not only on the candidates identified for sodium and molybdenum but also for the 22 other elements in both above and below ground tissues.

BIOGRAPHICAL SKETCH

Joshua Cobb was born on December 19, 1981 in rural Maine. Despite enduring what some would consider a disadvantaged childhood, he graduated high school 11th in his class at Erskine Academy. Following high school, he enrolled in Brigham Young University as microbiology major. After one semester he deferred his education for two years to accept a volunteer assignment to serve as a missionary for the Church of Jesus Christ of Latter-day Saints in Cordoba, Argentina. After his mission, Joshua resumed his studies at Brigham Young University and received a Bachelor's degree in Biotechnology in 2006 and continued on to receive a Master's degree in Genetics and Biotechnology in 2008. During which time he also met and married the love of his life and fathered their first two children. With his small family in tow, Joshua was accepted to the PhD program in Plant Breeding and Genetics at Cornell University. While at Cornell Joshua proposed and conducted an independent research project that coordinated efforts between three major internationally recognized laboratories and had opportunities to study international agriculture in Mexico, India, and the Philippines. He worked for three years as a TA for an introductory plant breeding course and was highly praised for his efforts to connect with and empower students. He and his wife also doubled their reproductive fitness with the addition of fraternal twins about 3 years prior to his graduation. His educational success even warranted an invitation from his high school alma mater to deliver the 2012 commencement address. Following his education at Cornell, Joshua has accepted a position with DuPont Pioneer as a molecular rice breeding scientist in Hyderabad, India. He hopes

to further develop his international professional network, gain valuable experience with field breeding programs, participate in the training and support of plant breeding and genetics programs in India and eventually return to the classroom where he can use his experience to inspire and train the next generation of agricultural scientists.

ACKNOWLEDGMENTS

It's hard for me to express how deeply grateful I am to the very special people who have been placed throughout my life and have enabled me to overcome seemingly insurmountable challenges. Especially to my family and friends who have always believed in me even when they had to empirical evidence for doing so. To my wife who for 10 years has endured the rigors and salary having a student-husband and my kids who consistently found opportunities to challenge me in ways that forced me to grow in directions I never could have otherwise.

For Susan whose clarity of thought and vision has never ceased to inspire me to become something more than myself. For her patience and dedication while I tried to meet my obligations to her, and to my family. She was a wellspring of wisdom and guidance in all areas of my life and I am a better person for having spent time under her tutelage. For Leon and the near immediate paternal relationship he helped cultivate with me. And the support and guidance he provided me. His cheerful disposition and selfless demeanor was my anchor when the storms of academic life raged around me.

I would also like to acknowledge the incredible people whose collaboration has made this ambitious project possible. Not the least of which are Jason Mezey, Anthony Greenberg, Pavel Korneliev, Genevieve DeClerck, Randy Clark, Ize Imai, Sam Leiboff, Jenny Cornell, Maria Carrizales, Michael Yang, and Rebecca Deveau-Greene

I also acknowledge the support of the world class faculty of the Cornell Plant Breeding and Genetics Department, not the least of which include Ronnie Coffman, Mark Sorrells, Tim Setter, and Margaret Smith each of which I feel a special connection to. I am grateful as well to the supportive and unified graduate student community within the department and the friendships I've forged as a result.

I also acknowledge and am grateful to the funding agencies that helped financially support my PhD, including fellowship support from DuPont Pioneer, a Cornell University Diversity Fellowship, the United States Department of Agriculture, and the National Science Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Chapter 2 was published as an invited review in Theoretical and Applied Genetics (Cobb et al. 2013, doi: 10.1007/s00122-013-2066-0) under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works as long as the author of the original work is cited. It includes three subsections initially authored by those other than the author of this dissertation and are as indicated in the manuscript.

Chapter 3 comprises the majority part of a paper to be submitted as an original research article to Nature Communications in mid-2013.

CHAPTER 1:

INTRODUCTION

The most important innovation of the last 13,000 years of human history is arguably the domestication of plants and animals. Plant domestication in particular was paramount to the formation of settlements and the migration from a nomadic hunter-gatherer lifestyle to settled, sedentary, and specialized communities. Relatively few wild plant species have undergone domestication and today more than 70% of the calories consumed by humans are provided by just 15 crops, with five (rice, wheat, maize, sugarcane, and barely) occupying more than fifty percent of that contribution (Ross-Ibarra et al. 2007). Today these crops must support more than 7 billion people and to meet that demand it's been estimated that the first 50 years of the 21st century the world will have to produce as much food as has been produced in the last 10,000 years (Hoisington et al. 1999).

The importance of rice and its genetic substructure

Domesticated rice (*Oryza sativa* L.) is the most important source of caloric intake for humans, serving as a staple for more than half of the global population and providing between 20% and 80% of dietary calories for most of Asia (Khush 1997). Additionally rice improvement is a good target for the alleviation of poverty and for international development efforts because Asia is home to 70% of the people living on less than a dollar a day and the poorest among them spend more than 50% of their income on rice (Zeigler and Barclay 2008). Increases in arable land under production alone cannot sustain the growing dependency on rice as a primary food source and as such a careful

consideration of the genetic heritage of domesticated plant species left to us by our Neolithic progenitors remains the only viable option to sustain growth throughout the 21st century.

It's widely hypothesized that rice domestication proceeded from two related but independently cultivated groups of germplasm that derived from pre-differentiated ancestral gene pools of *Oryza rufipogon*, a semi-perennial aquatic species of wild rice (Kovach and McCouch 2008; Nakano et al., 1992; Sun et al., 2002). Despite geographic and biological barriers, significant gene flow between the domesticated populations still occurred and was phenotypically impactful (Kovach et al. 2009; Londo et al., 2006; McCouch et al., 2012; Second, 1982). As rice and humans moved together through regions of ecological adaptation, eventually five distinct genetic subpopulations emerged in the two pre-differentiated clades (Garris et al. 2005). These five subpopulations form the basis of all modern breeding efforts and constrain the breeding objectives in different parts of the world. However, as the genetic separation between these gene pools is both ancient and wide, admixture between them provides unique and powerful opportunities to observe transgressive variation for important physiological and agronomic traits (Famoso et al 2011, Maas et al. 2010, Gutierrez et al. 2010, Septiningsih et al. 2003, Thomson et al. 2003, and Septiningsih et al. 2013).

The essentiality of mineral nutrients to plant life on earth

The processing and acquisition of minerals is essential to all life on earth. Many mineral elements are necessary for the proper functioning of biochemical processes

related to life and must be obtained through the diet. Because plants are autotrophic they have had to evolve mechanisms of absorbing all the appropriate mineral nutrients from the soil necessary to sustain life while at the same time avoiding the uptake of toxic minerals that sabotage physiological processes. Mineral elements essential for plant biology are defined as those that are necessary for a plant to complete its life cycle and for which no other element can substitute.

These essential elements are usually classified as either macronutrients or micronutrients, depending on their relative abundance within the organism. Macronutrients include nitrogen (N), potassium (K), calcium (Ca), magnesium (Mg), phosphorus (P), and sulfur (S) (Maathius 2009). Carbon, hydrogen, and oxygen are likewise abundant, but are derived from CO_2 and H_2O and are not usually treated as mineral nutrients. Essential plant micronutrients include boron (B), chlorine (Cl), copper (Cu), iron (Fe), manganese (Mn), molybdenum (Mo), nickel (Ni), and zinc (Zn) (Hansch and Mendel 2009). Additionally cobalt (Co), sodium (Na), selenium (Se) and silicon (Si) are non-essential but considered beneficial for certain plant species and in the right amounts and can promote plant growth in these species (Pilon-Smits et al. 2009). While each of these nutrients is important in its own sphere, there are also many toxic minerals that can highjack transporters for essential nutrients, find their way into the organism and disrupt important physiological processes that depend on other elements for which the toxins are chemical analogs (Nagajyoti et al. 2010).

All of these elements are required at different levels depending on the plant species in question, and almost any of them can be toxic if found in too great abundance (particularly the micronutrients). Soils with an overabundance of heavy metals or a

deficiency of nutrients pose pointed challenges for sessile organisms like plants in maintaining mineral nutrient homeostasis. Additionally a great variety of soil environmental conditions (pH, cation exchange capacity, electrical conductivity, water content etc.) can dramatically affect the availability of mineral nutrients in the soil. The application of nutrients and remediation of toxic metals in the soil are important in most agricultural settings, and in rice production specifically. That being the case, rice is also particularly interesting from the vantage point of mineral nutrient acquisition because it is grown in a wide variety of climates, soil chemistries, and agronomic conditions. There are rice varieties adapted to dry elevated upland hillsides, many of which have very acidic soils, rainfed lowland areas, semi-aquatic flooded conditions, deepwater conditions and mangrove swamps.

Resources and shifting limitations of rice genetic analysis

Rice is a convenient system to study because it holds the distinction of being both a major crop staple and a model organism. As such, there is great motivation in the scientific community to provide powerful tools for the genetic analysis of rice. Rice boasts two fully sequenced reference genomes (*temperate japonica* cv. Nipponbare and *indica* cv. 93-11). The former being a BAC-by-BAC sequencing approach and the later employing whole genome shot-gun re-sequencing supported by alignment to the Nipponbare sequence (Goff et al. 2002, Yu et al. 2002). The latest (MSU7) assembly of the Nipponbare reference genome notes 49,066 gene models, of which 39,045 are classified as genes (non-TE loci) (http://rice.plantbiology.msu.edu/). In addition to the reference genomes, there are also multitudes of mutagenized populations, whole-genome microarrays, genome tiling arrays (Jung et al. 2008) mapping populations (Ahmadi et al. 2005, Price et al. 1997, Mei et al. 2006, Tung et al. 2010, Lin et al.

1998, Bandillo et al. 2010), and gene co-expression network data (Sato et al. 2013) available to further the functional annotation of the rice genome and to associate phenotype with genotype. Interestingly, with the ever increasing availability and economy of these powerful tools, the primary constraint of genetic analysis has moved from the ability to collect sufficiently representative genomic information to being able to accurately phenotype appropriate germplasm in meaningful ways (See Chapter 2 of this dissertation).

This dissertation describes the nature of the 'phenotyping bottleneck' and demonstrates a powerful example of how multiple tools for genetic analysis can be combined with sophisticated phenotyping approaches to characterize natural variation for complex traits in rice.

Chapter 2 illustrates in detail the different applications of genotype-phenotype association analysis and discusses phenotyping concerns for each of them. Drought tolerance is presented as an example where evaluation under both controlled and field conditions contributes to the annotation of genomic regions and their phenotypic consequences. Additionally in collaboration with Randy Clark, Genevieve DeClerck, and Anthony Greenberg, we describe strategies to create modular and precise phenotyping platforms that empower the analysis rather than constrain it. We also highlight community efforts to database phenotypic information in a way that makes the data both accessible and query-able so as to motivate hypothesis driven research. Additionally advice is offered on statistical approaches to estimate phenotypic values for an individual and to integrate over the inherent variation in the measurements themselves. Lastly we conclude with a challenge to the community to come together

in distributed phenotyping networks across the globe through collaboration in order to leverage the comparative advantage of different ecological and geographical locations, and fully understand genotype by environment interaction.

In chapter 3, a novel and innovative phenotyping platform is described that takes an 'ionomics' approach to understanding the genetic architecture of nutrient acquisition in domesticated rice. The phenotyping platform accurately and precisely measures the concentration of 24 different mineral elements in rice roots and shoots in a diverse population of rice designed to represent all five domestic subpopulations. This is accomplished by utilizing a hydroponics platform and sophisticated analytical chemistry techniques. The phenotype data is then combined with high density and high quality genotype data in order to effectuate genome-wide association (GWA) mapping. The GWA information is combined with data generated from other research approaches (re-sequencing data from 125 genomes, candidate gene analysis using the Nipponbare reference genome, comparative genomics with *Arabidopsis thaliana* and *Zea mays*, rice gene co-expression network databases, and bi-parental introgression lines) to annotate peaks in the mapping analysis and draw hypothesis about the genetic architecture of both molybdenum and sodium shoot concentration. We identified two genes of major importance (*OsMOT1* and *OsSLT1*) for molybdenum and sodium shoot concentration respectively, and demonstrate the correlation between natural allelic variation at these loci and the phenotypic variance for each trait. Furthermore we model the phenotype data in a Bayesian framework in order to understand relationships between traits in the correlated phenotypic network that is the ionome. We highlight the strong negative correlation between shoot biomass and shoot cadmium concentration and demonstrate that this inverse relationship is specific to the *temperate japonica* and *indica* gene pools and that *aus* and *tropical japonica* are likely

targets for identifying alleles that regulate shoot cadmium accumulation even in the presence of the semi-dwarf phenotype.

In Chapter 4 several key areas are highlighted where a more in depth evaluation of ionomics phenotypes will continue along with an evaluation of the kinds of tools and analyses that will allow researchers to derive the most value from the large dataset we have generated. A bi-parental QTL analysis of the rice ionome that is currently underway on the Azucena x IR64 SSD RIL population is described as well as how QTL identified in this way may serve to validate GWA peaks. The pending development of RDP2 is also introduced, which is a much larger and more diverse panel of domesticated rice germplasm that will permit the detailed genetic interrogation of subpopulations in a statistically robust way. In addition to domesticated germplasm, we have collected ionomic data on 77 wild *O. rufipogon* genotypes which are destined for high density genotyping and will lend insight into the effects of domestication on the rice ionome as well as provide potentially powerful allelic variation that can be used to effect phenotypic modifications of agriculturally relevant germplasm. This will be best accomplished by using the chromosomal segment substitution lines (CSSLs) currently under development that are derived from wild donor genotypes. This will become especially important as this targeted germplasm can be effectively replicated and tested in the field. It is also important to emphasize that there is still a mountain of GWA data that remains as yet un-mined for a number of different elemental phenotypes and more yet to come as derivative phenotypes (such as root/shoot ratios and ratios between chemical analogs) are extracted from the phenotypic data. This becomes especially important as we continue to develop strategies for multivariate GWA analysis that consider localized linkage disequilibrium and estimate the significance of haplotypes rather than SNPs. Lastly,

the cell and molecular biology research that must be done to test the hypothesis generated as part of this thesis about the function and regulation of *OsMOT1* and *OsSLT1* is discussed.

REFERENCES

Ahmadi, N., C. Dubreuil-Tranchant, B. Courtois, D. Foncéka, D. This, S. R. McCouch, M. Lorieux, J. C. Glaszmann, and A. Ghesquière. "New resources and integrated maps for IR64× Azucena, a reference population in rice." In *IRRI 5th International Rice Genetics Symposium and 3rd International Rice Functional Genomics Symposium, Manila, Philip*, pp. 19-23. 2005.

Bandillo, N., P. A. Muyco, C. Caspillo, M. Laza, A. G. Sajise, R. K. Singh, G. B. Gregorio, E. Redona, and H. Leung. "Development of multiparent advanced generation intercross (magic) populations for gene discovery in rice (Oryza sativa L.)." *Philippine Journal of Crop Science* 35 (2010).

Famoso AN, Clark RT, Shaff JE, Craft E, McCouch SR, Kochian LV (2010) Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal aluminum tolerance and investigations into rice aluminum tolerance mechanisms. Plant Physiol 153:1678-1691

Garris AJ, Tai TH, Coburn J, Kresovich S, McCOUCH S (2005) Genetic structure and diversity in Oryza sativa L. Genetics 169:1631-1638

Goff SA, Ricke D, Lan T, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92-100

Gutierrez, A., Carabalí, S., Giraldo, O., Martínez, C., Correa, F., Prado, G., ... & Lorieux, M. (2010). Identification of a rice stripe necrosis virus resistance locus and yield component QTLs using Oryza sativa× O. glaberrima introgression lines. BMC plant biology, 10(1), 6.

Hänsch R, Mendel RR (2009) Physiological functions of mineral micronutrients (Cu, Zn, Mn, Fe, Ni, Mo, B, Cl). Curr Opin Plant Biol 12:259-266

Hoisington D, Khairallah M, Reeves T, Ribaut J, Skovmand B, Taba S, Warburton M (1999) Plant genetic resources: what can they contribute toward increased crop productivity?. Proceedings of the National Academy of Sciences 96:5937-5943

Khush GS (1997) Origin, dispersal, cultivation and variation of rice. In: Oryza: From Molecule to Plant. Springer, pp 25-34

Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch SR (2009) The origin and evolution of fragrance in rice (Oryza sativa L.). Proceedings of the National Academy of Sciences 106:14444-14449

Kovach M, McCouch S (2008) Leveraging natural diversity: back through the bottleneck. Curr Opin Plant Biol 11:193-200

Londo, J.P. et al. (2006) Phylogeography of Asian wild rice, Oryza rufipogon reveals multiple independent domestications of cultivated rice, Oryza sativa. Proc. Natl. Acad. Sci. U. S. A. 103, 9578–9583

Maas, L. F., McClung, A., & McCouch, S. (2010). Dissection of a QTL reveals an adaptive, interacting gene complex associated with transgressive variation for flowering time in rice. Theoretical and applied genetics, 120(5), 895-908.

Maathuis FJ (2009) Physiological functions of mineral macronutrients. Curr Opin Plant Biol 12:250-258

McCouch, SR, MJ Kovach, MT Sweeney, M Semon (2012) The dynamics of rice domestication: a balance between gene flow and genetic isolation. *In* Biodiversity in Agriculture. Domestication, Evolution, and Sustainability (eds.) Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE, Qualset CO. Cambridge University Press, New York. Chapter 13, Pp. 311-329

Nagajyoti P, Lee K, Sreekanth T (2010) Heavy metals, occurrence and toxicity for plants: a review. Environmental Chemistry Letters 8:199-216

Nakano, M. et al. (1992) Phylogenetic study of cultivated rice and its wild relatives by RFLP. Rice Genet. Nwsltr. 9, 132–134

Pilon-Smits EA, Quinn CF, Tapken W, Malagoli M, Schiavon M (2009) Physiological functions of beneficial elements. Curr Opin Plant Biol 12:267-274

Ross-Ibarra J, Morrell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proceedings of the National Academy of Sciences 104:8641-8648

Sato Y, Namiki N, Takehisa H, Kamatsuki K, Minami H, Ikawa H, Ohyanagi H, Sugimoto K, Itoh J, Antonio B, Nagamura Y (2013) RiceFREND: a platform for retrieving coexpressed gene networks in rice
Nucleic Acids Research 41:D1214-D1221

Second, G. (1982) Origin of the genetic diversity of cultivated rice (*Oryza* spp.): study of the polymorphism scored at 40 isozyme loci. Jpn. J. Genet. 57, 25–57

Septiningsih, E. M., Ignacio, J. C. I., Sendon, P. M., Sanchez, D. L., Ismail, A. M., & Mackill, D. J. (2013). QTL mapping and confirmation for tolerance of anaerobic conditions during germination derived from the rice landrace Ma-Zhan Red. Theoretical and Applied Genetics, 1-10.

Septiningsih EM, Prasetiyono J, Lubis E, Tai TH, Tjubaryat T, Moeljopawiro S, McCouch SR (2003) Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the Oryza sativa variety IR64 and the wild relative O. rufipogon. Theor Appl Genet 107:1419–1432

Sun, Q. et al. (2002) Genetic differentiation for nuclear mitochondrial and chloroplast genomes in common wild rice (O. rufipogon Griff) and cultivated rice (*O. sativa* L). Theor. Appl. Genet. 104, 1335–1345

Tung C, Zhao K, Wright MH, Ali ML, Jung J, Kimball J, Tyagi W, Thomson MJ, McNally K, Leung H, Kim H, Ahn S, Reynolds A, Scheffler B, Eizenga G, McClung A, Bustamante C, McCouch SR (2010) Development of a Research Platform for Dissecting Phenotype–Genotype Associations in Rice (Oryza spp.). Rice. doi: 10.1007/s12284-010-9056-5

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296:79-92

Zeigler RS, Barclay A (2008) The relevance of rice. Rice 1:3-10

CHAPTER 2:

NEXT GENERATION PHENOTYPING: REQUIREMENTS AND STRATEGIES FOR ENHANCING OUR UNDERSTANDING OF GENOTYPE-PHENOTYPE RELATIONSHIPS AND ITS RELEVANCE TO CROP IMPROVEMENT

Abstract

More accurate and precise phenotyping strategies are necessary to empower high-resolution linkage mapping and genome wide association studies (GWAS) and for training genomic selection (GS) models in plant improvement. Within this framework, the objective of modern phenotyping is to increase the accuracy, precision and throughput of phenotypic estimation at all levels of biological organization while reducing costs and minimizing labor through automation, remote sensing, improved data integration and experimental design. Much like the efforts to optimize genotyping during the 1980's and 1990's, designing effective phenotyping initiatives today requires multi-faceted collaborations between biologists, computer scientists, statisticians and engineers. Robust phenotyping systems are needed to characterize the full suite of genetic factors that contribute to quantitative phenotypic variation across cells, organs and tissues, developmental stages, years, environments, species and research programs. Next generation phenotyping generates significantly more data than previously and requires novel data management, access and storage systems, increased use of ontologies to facilitate data integration, and new statistical tools for enhancing experimental design and extracting biologically meaningful signal from

environmental and experimental noise. To ensure relevance, the implementation of efficient and informative phenotyping experiments also requires familiarity with diverse germplasm resources, population structures, and target populations of environments. Today, phenotyping is quickly rising as the major operational bottleneck limiting the power of genetic analysis and genomic prediction. The challenge for the next generation of quantitative geneticists and plant breeders is not only to understand the genetic basis of complex trait variation but also to use that knowledge to efficiently synthesize 21st century crop varieties.

Key Words

High-throughput Phenotyping, Genomic Selection, Genome-wide Association, Database Management, Bayesian Analysis

Introduction

Agriculture faces tremendous challenges in the decades ahead. The FAO predicts that population and income growth will double the global demand for food by 2050, effectively increasing competition for crops as sources of bioenergy, fiber and for other industrial purposes (http://www.fao.org). Compounding the pressure for increased agricultural output are looming threats of water scarcity, soil fertility constraints, and climate change. Addressing these problems will require innovative approaches to both the agronomic and the genetic components of crop production systems. More sustainable management of renewable soil and water resources, in concert with more efficient utilization of genetic diversity will be key to achieving the

necessary gains in productivity (Bakker et al. 2012, Frison et al. 2011, Cai et al. 2011, Pypers et al. 2011, McCouch et al. 2012).

Genetic diversity provides the basis for all plant improvement. Historically, plant breeders have sought to understand the nature of genetic variation by evaluating the performance of breeding populations over years and locations. Using replication and sophisticated experimental designs, they obtained useful insights about trait heritability, the influence of environment, the breeding value of different parents, and strategies for selecting genetically superior offspring in the field. With the dawn of the genomics era, emphasis began to shift toward the evaluation of genetic diversity directly at the DNA level. This approach is of interest to geneticists for the evolutionary and functional insights it brings, and to plant breeders as a source of tools for improving the power and efficiency of selection. Parallel investments in genotyping and phenotyping have generated datasets that can be associated with each other to address both basic and applied questions. Geneticists are interested in the nature and origin of mutations and their functional significance in the context of both qualitative and quantitative traits. Plant breeders embrace genomics as a way to document and protect the genetic composition of plant varieties, trace pedigree relationships, identify and select valuable mutations, and gain insight into the nature of genotype by genotype (GXG) and genotype by environment (GXE) interactions. The ultimate goal of genomics research in plant breeding is to contribute to improving the efficiency, effectiveness and economy of cultivar improvement.

As biology moves from a data-starved and largely observational discipline to a data-rich science capable of prediction, it follows the path of physics and engineering that came before. The tremendous innovation in genomics technology over the last two decades has been driven by multi-faceted collaborations between chemists, biologists and engineers, and today, costs continue to decline while accuracy and throughput continue to increase (Elshire et al. 2011, Tung et al. 2010). Correlated with the downward trend in the cost of sequencing is the expanded use of high-resolution genotyping in plant species that were previously ignored by the genomics community, a sampling of which include cassava, common bean, pea, sunflower, cowpea, and grain amaranth (Bachlava et al. 2012, Ferguson et al. 2011, Hyten et al. 2010, Maughan et al. 2011, Smýkal et al. 2012, Varshney et al. 2009, Varshney et al. 2010). In addition to offering new insights into diverse germplasm resources, high throughput genotyping and next-generation sequencing (NGS) make it possible to efficiently leverage genetic information across species. The power of whole-genome sequencing as a unifying force in biology has motivated the development of diversity panels and large mapping populations in many crop species to facilitate trait dissection and gene discovery (Atwell et al. 2010, Huang et al. 2010, McCouch et al. 2012, Yu et al. 2008, Zhao et al. 2011a, Neumann et al. 2011, Pasam et al. 2012). It has also catalyzed new thinking about how to manipulate the genetic variation that exists in elite gene pools (Chen et al. 2011, Thomson et al. 2011, Trebbi et al. 2011).

With the deluge of low-cost genomic information on important crop species, a fundamental change in research emphasis is needed to address the shortage of high

quality phenotypic information. At this time, phenotyping has replaced genotyping as the major operational bottleneck and funding constraint of genetic analyses. Unlike genotyping, which is now highly mechanized and essentially uniform across organisms, phenotyping is still a cottage industry, species-specific, labor intensive, and inevitably environmentally sensitive. Further, while sequence variation is theoretically finite, and thus all sequence variants could conceivably be discovered for a given crop species, there is no expectation that the phenome will ever be fully characterized (Houle et al. 2010). The phenome of an organism is dynamic and conditional, representing a complex set of responses to a multi-dimensional set of endogenous and exogenous signals that are integrated over the evolutionary and developmental life history of an individual. Phenotypic information can be envisioned as a continuous stream of data that changes over the course of development of species, a population or an individual in response to different environmental conditions. It can be associated with genomic information to understand the components of phenotypic variation that are due to genetics. With increasing availability of high-density genotypic information, understanding genotype-phenotype relationships becomes increasingly dependent on the availability of high quality phenotypic and environmental information.

Over the next two decades, the development of phenotyping strategies will almost certainly mirror innovations in genotyping technology that have occurred over the last twenty years, characterized by increasing automation and throughput (Rafalski and Tingey 1993, Perlin et al. 1995, Sheffield et al. 1995, Weber and Broman 2001). As

the science of phenotyping evolves, emphasis will increasingly be placed on generating information that is as accurate (able to effectively measure traits and/or performance characteristics), precise (small variance associated with replicated measurement), and as relevant as possible, while keeping costs within reasonable limits. If developments in genotyping offer a roadmap for where phenotyping is going in the future, these objectives will be reached based on new forms of automation and collaborations between biologists, engineers and computer scientists.

The purpose of this review is to outline considerations related to the future of phenotyping as the basis for association mapping and gene discovery as well as for developing predictive genomic selection models for crop improvement.

Association between Phenotype and Genotype

The central challenge of modern genetic analysis is to understand the biological determinants of quantitative phenotypic variation. To date efforts in the plant genetics community have done well at identifying genes underlying traits controlled by one or a few loci with large effects. This is particularly true in the major crop species where genetic analyses have identified the biochemical basis of many important phenotypes (particularly resistance to biotic and abiotic stress) and have also been the driving force behind the development of tools for marker assisted selection in crop improvement (Foolad and Panthee 2012, Jin et al. 2010, Paux et al. 2010, Robbins et al. 2010). However, understanding complex trait variation has proven frustratingly difficult, as the genetic architecture of these important traits often involves many loci

of small effect that may interact with each other as well as with the environment (Buckler et al. 2009, Collard and Mackill 2008, Schuster 2011). To discriminate such small effects, a combination of technologies and statistical methods are now being employed. Next-generation sequencing technologies have provided an economically feasible way to survey genetic variation with a resolution that is now limited more by the linkage disequilibrium (LD) in a particular mapping population than by marker density. This phenomenon has motivated the assembly of large panels of genetic diversity as well as the creation of large inter-mated populations to manipulate LD and facilitate the association of genotype with phenotype (Huang et al. 2011, Yu et al. 2008, Zhao et al. 2011a). These large and diverse populations aim to increase the recombination frequency and the frequency of rare alleles in order to enhance the power to infer the effects of individual loci. This also highlights the need for careful population design and advocates for the inclusion of admixed lines that may provide statistically useful observations of allele effects in diverse genetic backgrounds.

Phenotyping for Genomic Selection

The emphasis on precision-phenotyping represents a significant change for breeders engaged in variety development who have traditionally favored simplicity, speed, and flexibility over sensitivity, precision and accuracy. This is because historically, the advantages of the latter could not be translated into economically relevant genetic gain in a breeding context. We argue that this paradigm is beginning to change with the potential to integrate genomic selection into a variety development program. As the cost and efficiency of obtaining genomic information on large numbers of individuals

dips below the cost and efficiency of evaluating populations phenotypically over years and environments, the breeding community is alert to the idea that genomic information can be leveraged to predict phenotypic outcomes (Cabrera-Bosquet et al. 2012, Heffner et al. 2009, Heslot et al. 2012). Further, the use of Bayesian models facilitates the analysis of sparse data (where not all individuals or families are evaluated phenotypically in each environment) and strongly suggests that there are cost-effective experimental designs that can dramatically reduce the amount of replication needed to extract meaningful phenotypic performance indicators for a population (see section on "Analysis, Adjustment, and Value extraction of Phenotypic Data").

If the accuracy of genomic predictions is sufficient to offset the time and expense required to evaluate the performance of the breeding populations in the traditional manner, and if genomic selection demonstrates a clear increase in the rate of genetic gain per cycle of selection, then breeders will quickly adopt the most efficient strategy to accomplish their goals. This may require staggered use of traditional and precision-phenotyping, depending on the trait(s) and the species under consideration. What is important is breeders begin to reevaluate how a focused investment in precision-phenotyping of a training population may be able to minimize the requirement for costly, extensive phenotyping of large numbers of lines every generation in the future. The purpose of this paper is to explore some of the key dimensions of next-generation phenotyping that will allow geneticists and breeders to productively interrogate the

complex ménage-à-trois between genotype, phenotype and the environment and develop models to leverage genotypic information to predict phenotypic outcomes.

Under a genomic selection model, precision phenotyping is most important when evaluating a training population because that dataset provides the basis for developing the statistical model that is then used to predict phenotypic performance in related members of a breeding population. The model is derived from the relationship between phenotype, genotype, and GXE, where marker genotypes are treated as random variables. Genomic selection is particularly useful when it can save a generation or two of time-consuming and expensive phenotyping, as only comparatively small training populations need be screened.

Genomic selection aims to model genome-wide SNP variation without concern for identifying particular alleles, loci or pathways or understanding how different alleles contribute to the phenotype. Since the metric of success is the ability to predict the performance of an adapted line or variety under relevant agronomic conditions, it is important to consider phenotyping strategies that 1) estimate crop performance under appropriate management conditions in the field, 2) can evaluate performance across a population of target environments, and 3) can generate useful data in real time without a disproportionate investment in labor and infrastructure. Despite the advantages of accelerating the breeding cycle, the ability of genomic selection models to accurately predict phenotypes on is dependent on using prohibitively large training populations when working with traits with low heritability and complex inheritance (Calus et al.

2008, Guo et al. 2012, Hayes et al. 2009, Heffner et al. 2010, Jia and Jannink 2012, Kumar et al. 2012, Nakaya and Isobe 2012, Munoz et al. 2011, Resende Jr et al. 2012, Zhao et al. 2011b, Zhong et al. 2009). This is due to the fact that GXE interaction plays a major role in explaining field performance, and genomic selection is highly dependent on a prediction model developed from a limited sampling of the environmental variance. Further research is needed to improve the accuracy of prediction under a genomic selection model, particularly in the face of climate change and increased environmental variance across the target population of environments.

Phenotyping for QTL and gene discovery

In contrast to genomic selection, phenotyping of a diversity panel for Genome Wide Association Studies (GWAS) or a bi-parental mapping population for QTL analysis is designed to interpret and dissect the genetic architecture of complex traits and to understand how specific DNA variants condition the inheritance of diverse phenotypes. Both forms of linkage mapping are successful at implicating genomic regions involved in complex trait variation, but cloning the gene(s) underlying the QTL remains time consuming and resource intensive, even when the QTL explains a substantial proportion of the phenotypic variation (Bhattacharyya 2010, Fan et al. 2006, Krattinger et al. 2009, Li et al. 2010, Liu et al. 2008, Saito et al. 2010). Bi-parental populations are limited by the particular alleles present in the parents but they offer power for QTL dissection because population structure is disrupted and genetic background differences in the progeny are constrained. Association mapping studies, on the other hand, generally provide higher resolution of QTL for the same number of

lines and evaluate a wider array of alleles but are limited by the inability to interrogate rare alleles or to dissect phenotypes that are perfectly correlated with population structure (Kruglyak, 2008; Manolio and Collins, 2009; Price et al., 2006; Pritchard and Cox, 2002; Reich and Lander, 2001). When large numbers of markers are used for either QTL analysis or GWAS, a multiple test correction is required to limit the false discovery rate . With ever-improving statistical approaches and improvements in the accuracy and precision of phenotyping, both forms of linkage mapping hold great promise for elucidating the genetic architecture of complex traits and identifying the genes and specific alleles underlying trait variation.

Sampling vs. Controlling Environmental Variation

Different approaches to phenotyping are required for different purposes (Campos et al. 2004, Crouch et al. 2009, Gordon and Finch 2005, Kloth et al. 2012, Masuka et al. 2012, Pieruschka and Poorter 2012). Plant breeders have traditionally relied on large-scale replication of phenotypic trials over years and locations to identify individual families or populations that perform best in a target population of environments. By modeling locations and years as random effects, they were able to reliably extract genetic signal from environmental noise and identify varieties with broad or narrow zones of adaptation (Beavis 1998), though the process was very time and labor consuming. Many geneticists, on the other hand study phenotypic variation at the cell or tissue-specific level using plants grown under carefully defined environmental conditions, and evaluate cascades of molecular events using biochemical and “omics” technology. The world of the plant breeder and that of the molecular geneticist

intersect at the level of the plant, but the different scales of phenotyping make it challenging to integrate the knowledge contributed by each community into a unified and comprehensive view of the genetic determinants of plant growth, development and response to environment.

Under field conditions, it is often convenient to collapse quantitative phenotypes into discrete categories to facilitate manual data collection in real time and at reasonable cost. This has been the practice for many years among breeders and geneticists working with large, field-grown populations, and different communities of researchers have developed standardized categorical scales or indices for important whole-plant phenotypes that are easy to apply (Clarke et al. 1992, De Boever et al. 1993, International Rice Research Institute 1996, Kuhn and Smith 1977, Molina-Cano 1987, Yuan et al. 2004). For example, traits such as flowering time or disease resistance are frequently estimated using a visual assessment of "days to 50% flowering" in a row or plot, or "diseased leaf area" on individual plants. Historically, trait evaluation using these indices was reliable enough to provide reasonable data in the context of plant breeding. However, new population designs (Yu et al. 2008) in combination with high-density marker coverage have increased the power to detect small-effect QTL and estimate their effects, even on whole-plant phenotypes. This suggests that more rigorous, quantitative approaches to phenotyping are likely to bring rewards. Further, when there is significant variability in phenotypic scores collected by different individuals, more objective phenotyping protocols are desirable (Poland and Nelson 2010).

Recently, it has been argued that automated, high-throughput, field-based organismal phenotyping techniques involving remote sensing, such as near-infrared spectroscopy mounted on agricultural harvesters to measure spectral canopy reflectance with the aid of global positioning system (GPS)-guided tractors will enhance the precision and accuracy of phenotyping without extracting plants from the production environment (Cabrera-Bosquet et al. 2012, Houle et al. 2010, Montes et al. 2007, Tuberosa 2012, White et al. 2012). While these efforts can certainly facilitate selection for enhanced performance in a target zone of adaptation, one of the biggest challenges associated with these automated, field-based technologies is the variable nature of most natural environments.

To enhance the ability to screen for stress tolerance in field-grown plants, scientists often use plant populations to 'sample' the degree of stress encountered in a target population of environments (TPE). Once this has been ascertained, the TPE is used to evaluate the relative performance of different populations over several growing seasons. This requires significant up-front investment, as many different locations must be tested over multiple years in order to make an accurate estimation. Alternatively, breeders use “managed stress” as a way of optimizing screening protocols for application to large plant populations in the field. By managing the amount and timing of water, fertilizer, pest control or soil amendments, plants can be exposed to fairly reliable levels of stress while experiencing normal temperature, day length, etc. over the course of the growing season. These approaches work well if the

genetic component of phenotypic variation (heritability) is high, and if the differences among populations or individuals within a population are large. However, in cases where complex traits are conditioned by many alleles with small effects, the error associated with estimating the phenotype and the environmental variance contributing to the observed phenotypic variation are likely to dilute the relatively weak genetic signals and may preclude their detection.

To partially overcome this problem, many researchers have endeavored to take advantage of phenotyping strategies based on analytical chemistry (i.e. gas chromatography-mass spectroscopy, high performance liquid chromatography, inductively coupled plasma spectroscopy etc.) or a wide range of -omics technologies (transcriptomics, metabolomics, ionomics, proteomics, etc.). These are all highly automated and are important and useful due to their high-throughput and high accuracy. They are generally used to analyze specific anatomical parts of a plant at a particular time(s) in its development, and are best used on plants grown under well-defined growing conditions. Owing to the high cost per sample and the requirement for considerable technical expertise and infrastructure, these techniques may not be available to everyone and it may not be economically feasible to survey large numbers of field-grown plants. Thus, it often makes sense to first screen a population under controlled conditions with minimal replication and once a hypothesis about the genetic control of a trait of interest is formulated, it can be tested in a focused way in the field, or simply used to eliminate a large proportion of a population prior to undertaking field evaluation.

Screening populations under controlled conditions is also appropriate when the controlled environment is necessary to impose a particular form of stress or to permit growth of plants under specific conditions that cannot be replicated in the field. Controlled environments have been successfully used to inoculate plants with a particular strain of a pathogen, or to impose a particular abiotic stress such as aluminum toxicity without the natural coupling with phosphorus deficiency, or high CO_2 in combination with a critical night time temperature. Use of multi-step strategies involving both controlled and field environments are often the best way to maximize the extraction of useful genetic information while minimizing the expense and time involved (Fernie and Schauer 2009, Rafalski 2010).

Drought tolerance as an illustration of next-generation

While a complete survey of advances in drought phenotyping is beyond the scope of this review (See Mir et al. 2012 for a detailed overview of this topic), drought tolerance offers a compelling example of a combined approach of leveraging both controlled and uncontrolled phenotyping designs to enhance genetic analysis. The onset of water deficit and its impact on plant performance is a dynamic process that occurs across space and time. Under field conditions the inability to obtain standardized and consistent drought stress contributes to a loss in heritability and presents a challenge for both selection and mapping experiments (Berger et al. 2010). Many different approaches have been used to apply defined levels of drought stress in an effort to understand the nature of this complex trait, ranging from chemically

manipulating osmotic balance in hydroponics (Rengasamy 2010, Tavakkoli et al. 2010) to conveyer systems in glasshouses with digitally controlled irrigation systems (Granier et al. 2006, Jansen et al. 2009, Neumann 2013, Pereyra-Irujo et al. 2012) to the use of rainout shelters in the field (Czyczyło-Mysza et al. 2011, Dodig et al. 2012, Zhu et al. 2011a). Measurements of drought tolerance likewise range from surveys of root system architecture (Ibrahim et al. 2012, Landi et al. 2010, Lopes et al. 2011, Steele et al. 2007, Zhu et al. 2011b, Clark et al., 2011) to physiological metrics related to water status (Bartlett et al. 2012a, Bartlett et al. 2012b, Blum 2009, Gilbert et al. 2011, Ogburn and Edwards 2012, Tucker et al. 2011) to spectral imaging of shoot tissue (Berger et al. 2010, Goltsev et al. 2012, Liu et al. 2011, Zia et al. 2012) to simply evaluating yield under stress in the field (Bennett et al. 2012, Bernier et al. 2007, Ghimire et al. 2012, Golabadi et al. 2011, Messmer et al. 2009, Rehman et al. 2011, Swamy et al. 2011, Venuprasad et al. 2012, Vikram et al. 2011). Screening can be done using in-house facilities (growth chambers, green houses) or outsourced to a phenotyping facility such as the Jülich Plant Phenotyping Centre- JPPC (Jülich , Germany; http://www.fz-juelich.de/ibg/ibg-2/EN/About_us/organisation/JPPC/JPPC_node.html9), the Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung – IPK (Gatersleben, Germany; http://www.ipk-gatersleben.de/), the Plant Accelerator ® (University of Adelaide, Australia; http://www.plantaccelerator.org.au/), or the High Resolution Plant Phenomics Center –HRPPC (CSIRO Plant Industry, Canberra Australia; http://www.plantphenomics.org/HRPPC). The two former facilities are part of the larger European Plant Phenomics Network (http://www.plant-phenotyping-

network.eu/) and the latter two are part of the Australian Plant Phenomics Facility (http://www.plantphenomics.org.au/). Each system presents it's own advantages and disadvantages, but collectively they empower the researcher to investigate plant response to drought in ways that are more comprehensive than any one design can offer. These approaches are most often utilized for linkage mapping and gene discovery, and once QTL or candidate genes are identified, they can be validated for practical application by evaluating specific germplasm, genetic stocks or breeding populations under managed drought conditions in the field (Ali et al. 2010, Cavanagh et al. 2008, Huang et al. 2012, Kholová et al. 2010, Saisho and Takeda 2011, Venuprasad et al. 2011, Yadav et al. 2011). If validated, lines carrying the genes or QTLs of interest will be useful for elucidating the molecular mechanism(s) involved in the component trait(s), and will also be of immediate value as donor material for breeding with elite germplasm.

Examining the relationship between phenotypic variation under controlled environments and that observed under field conditions offers valuable insights that can be used to iteratively improve controlled environment phenotyping techniques so they are more predictive of plant performance in the field (Figure 2.1; Table 2.1). Ultimately, the choice of phenotyping approach will depend on the intention of the researcher, the size of the population in question (e.g. less than 10 individuals for precise physiological experiments, to a moderate number of lines (200-400) for mapping studies or GS training populations, or a large number of lines (400 – 1000+)

for association studies), the heritability of the phenotype, the tractability of the phenotype to controlled environment testing, and resource availability.

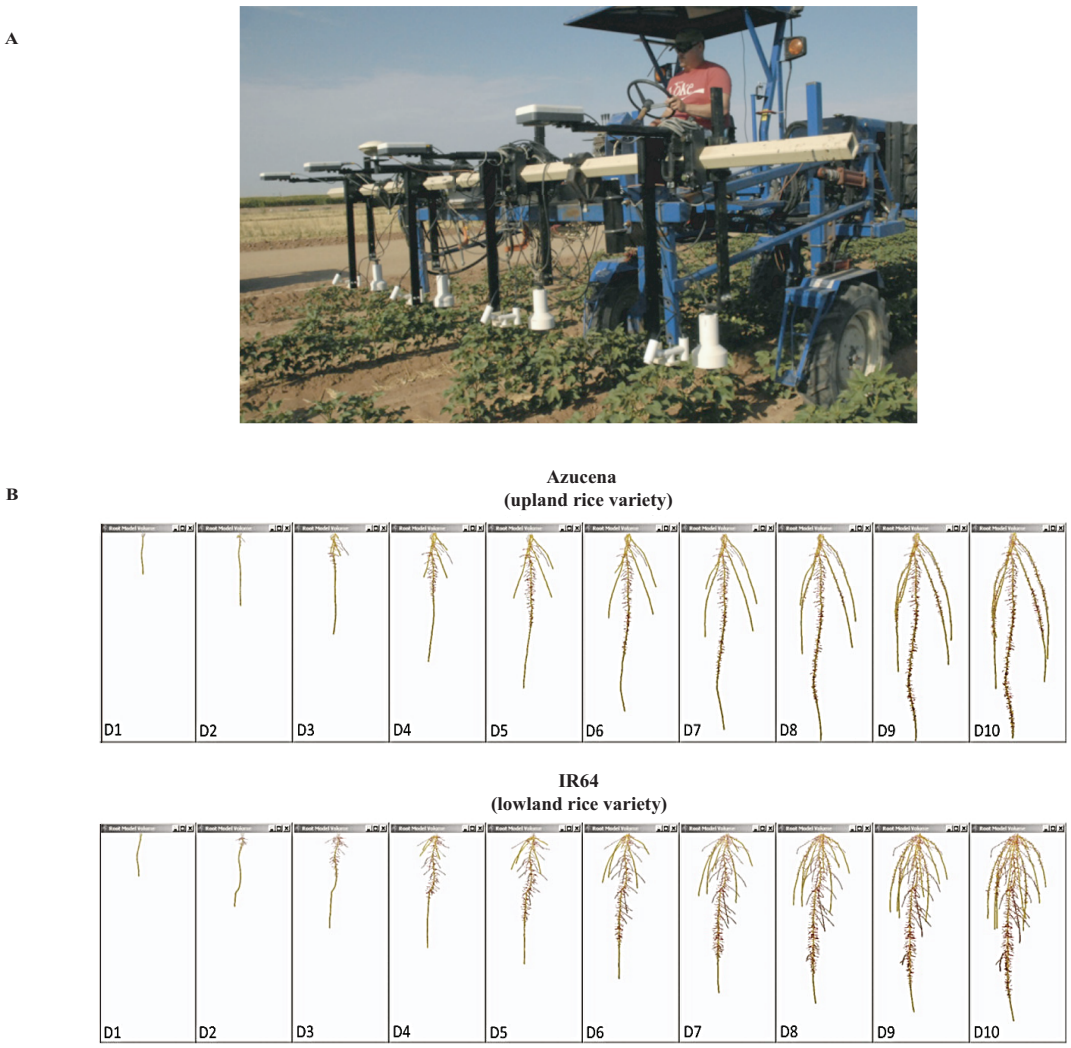


Figure 2.1:Illustration of different phenotyping platforms. The choice of phenotyping under controlled vs. field conditions depends on the purpose of phenotyping, heritability, and logistical considerations. A) High clearance tractor measuring the height, temperature, and spectral reflectance of young cotton plants from White et al. 2012. Such a system is high throughput and can measure canopy traits with high accuracy and precision. These traits have high heritability and are component phenotypes of yield under drought stress (*Reprinted from White et. al. Copyright 2012, with permission from Elsevier*). B) 10-day time course of root system growth in three dimensions of two divergent varieties of rice from Clark et al. 2012. Roots are notoriously difficult to phenotype in the field. This phenotype lends itself well to controlled conditions as the logistics of evaluating roots are more tractable, and it permits the exploration of otherwise un-surveyable phenotypes such as center of mass and dynamic tracking of architecture development over time (*copyright American Society of Plant Biologists*).

Controlled conditions	Field conditions
Minimizes environmental variation and increases heritability	Maximizes relevance to breeders and farmers
Increased precision of critical measurements	Characterizes the range of environmental variation
Maximizes information from a minimum of replicates	Evaluates over time as well as space
Decreases cost through automation and standardization	Estimates genotype × environment interaction
Develop hypotheses to be tested on a targeted set of lines in the field	Refine hypothesis and develop new screening protocols

Table 2.1. Controlled vs. Uncontrolled environments. Factors to consider when evaluating if a field based or environmentally controlled phenotyping platform is most appropriate.

Development of Technology and Phenotyping Tools (This sub-section initially authored by Dr. Randy Clark, Ph.D.)

The creative use of technology and careful development of tools to automate processes without sacrificing predictive power will be critical as next-generation phenotyping platforms are developed. This can be a real challenge as many experimental techniques in plant physiology, molecular biology and breeding can be restrictive and require specialized protocols that are often difficult to standardize. The integration of these approaches will be necessary to fully interrogate the genetic landscape of complex traits. Standardized phenotyping systems are not feasible for all research questions, but with thorough consideration and clearly defined objectives, many

techniques can be harnessed to investigate specific traits under high-throughput settings.

In recent years, automation, imaging, and software solutions have paved the way for many high-throughput phenotyping studies. Semi-automated systems have been successfully applied to investigate various components of plant growth and development, and can be used to help tackle basic research questions when combined with genetic mapping strategies (Famoso et al. 2010). Additionally, automated systems have allowed researchers to reduce the labor needed to manage and perform large-scale growth screens in laboratory, greenhouse and field environments (Nagel et al. 2012).

Aside from mechanization, digital imaging has emerged as a cornerstone to capturing quantitative phenotypic information under most automated or semi-automated approaches. Imaging has allowed many aspects of plant development, function, and health to be monitored, measured and tracked in ways previously unattainable using conventional metrics. Large image data sets, however, require novel software solutions in order to process and extract meaningful estimates of phenotypic variation. Most image analysis tools for plant phenotyping incorporate predefined processing and analysis procedures into semi-automatic or automatic routines in order to quantify multiple phenotypes from single images or groups of images.

In its essence, high-throughput phenotyping means integrating and optimizing a phenotyping process in a way that makes it as efficient and controlled as possible. In considering efficiency, several questions and decisions arise related to the accuracy, precision, automation, and adaptability of various stages of the phenotyping process, from growth techniques to experimental design and management practices to data capture and analysis strategies.

The accuracy and precision of the treatment and measurement process is a fundamental concern during any experimental procedure. During phenotyping studies where multiple individuals and replicates from different genetic backgrounds are evaluated across batches, effectively controlling the accuracy and precision of the phenotyping system will have direct impacts on the outcome of the analysis. Accuracy and precision are intimately interrelated, where accuracy represents how close the process or measurement is to the absolute truth and precision represents the repeatability or variance of the measurement process. Accuracy is important when there is variation across individual genotypes during mapping experiments. For instance, Clark et al. (2011) characterized the root systems of a rice bi-parental recombinant inbred line (RIL) population, and found that one parental genotype had dense, highly branched root systems while the other had long, sparse root systems. In order to clearly capture these differences a system needed to be designed that could correctly quantify both types of root systems in order to properly assess their relationship and further analyze variation within the progeny. Comparison and validation studies with known standards, such as the use of complementary imaging

modalities or other quantification software, can help evaluate the accuracy of a system.

Precision is critical when individual genotypes have multiple replicates that are evaluated across several batches. Presuming that the replicates share similar characteristics, the measurement system must be able to quantify the features in a repeatable way to prevent unpredictable system noise from masking the true similarities/differences between the genotypes. While there are statistical approaches for accounting for unwarranted variability *in silico*, efforts to improve the precision of data collection will only serve to enhance the statistical power of any analysis performed. The key to maintaining precision throughout a phenotyping activity is to employ stable instrument designs that can effectively control precision, such as the fixed lighting and camera setups used in the root systems in studies cited previously.

Unfortunately, there will always be tradeoffs between the maintenance of accuracy, precision, and the ultimate throughput of the phenotyping approach. As throughput and standardization increase, it necessitates a drop in accuracy and precision that must be carefully monitored in order to maintain the economic feasibility of the data collection. It is not always straightforward how to properly balance these tradeoffs, but through iterative design and testing, phenotyping tools can be established to satisfy research objectives and meet resource constraints.

The level of automation employed by a phenotyping approach is counter-balanced not only by tradeoffs with accuracy and precision, but also with adaptability. Increasing automation improves throughput and reduces labor costs, but also results in more specialized designs that have less adaptability and predictive power, and are prone to errors from non-standard individuals. This principle is illustrated well when image analysis involves batch processing many photographs using predefined algorithms and commands. It is fairly obvious that batch processing is invaluable during large-scale phenotyping experiments where thousands of images can be generated daily, but this also means that the software must rely on a rigid set of constraints. The quality of the images is usually not a problem during high-throughput phenotyping where the imaging process is standardized, but if any individuals deviate from pre-specified growth assumptions of the measurement algorithm, unpredictable and misleading measurement errors can arise. Even with automated analysis algorithms that have been thoroughly tested, it is necessary for the experimenter to manually check and validate the system outputs regularly. Along those same lines, incorporating user-guided processes into the phenotyping pipeline can also provide a useful compromise that improves the flexibility while maintaining the efficiency needed to perform large experiments (Clark et al. 2012, Le Bot et al. 2010, Lobet et al. 2011).

Most phenotyping tools that have been developed by research groups in the public and private sector are integrated in a way that makes them easy to disseminate and use, but sometimes this convenience can limit the range of their functionality to other studies. While this has precipitated the release of a number of software programs available for

the extraction of phenotype data from images (Table 2.2) the highly specific nature of individual phenotypes also motivates the development of in-house tools ideally suited to the analysis at hand. Although it is not a simple task, implementing modular designs will help increase flexibility of phenotyping in the future. The ImageJ analysis tool is a good example of the successful incorporation of modular designs in the software realm (Schneider et al. 2012). This image processing software allows users to create and share custom-developed plugins that expand the functionality of the software and make it applicable to wide range of research disciplines. Modular concepts have proven quite successful for the high-throughput phenotyping of notoriously difficult phenotypes such as root system traits. A notable example is GiaRoots, a software program that allows users to incorporate their own processing and algorithms into the automated analysis routines as a way of overcoming the limitations imposed by more integrated approaches (Galkovskyi et al. 2012).

Tissue	Software	Purpose and design	Reference
Roots	WinRhizo Tron	Morphological descriptions of root area, volume, length, etc	http://www.regent.qc.ca/products/rhizo/RHIZOTron.html
	KineRoot	2D analysis of root growth and curvature	Basu et al. (2007)
	PlaRoM	Platform for measuring root extension and growth traits	Yazdanbakhsh and Fisahn (2009)
	EZ-Rhizo	2D analysis of root system architecture	Armengaud et al. (2009)
	RoofTrace	Counting and measuring root morphology	Naeem et al. (2011), French et al. (2009)
	DART	2D analysis of root system architecture	Le Bot et al. (2010)
	SmartRoot	ImageJ plugin for the quantification of growth and architecture	Lobet et al. (2011)
	RootReader3D	3D analysis of root system architecture	Clark et al. (2011)
	RootReader2D	2D analysis of root system architecture	Clark et al. (2012)
	Gia-Roots	2D analysis of root system architecture	Galkovskyi et al. (2012)
Shoots/leaves	WinFolia	Morphological measurements of broad leaves	http://www.regent.qc.ca/products/folia/WinFOLIA.html
	TraitMill	Platform for measuring various agronomic characteristics	Reuzeau et al. (2006)
	PHENOPSIS	Automated measurement of water deficit-related traits	Granier et al. (2006)
	LeafAnalyser	Rapid analysis of leaf shape variation	Weight et al. (2007)
	LAMINA	Quantification of leaf size and shape	Bylesjö et al. (2008)
	HYPOTrace	Analysis if hypocotyl growth and shape	Wang et al. (2009)
	LEAFPROCESSOR	Analysis of leaf shape	Backhaus et al. (2010)
	Lamina2Shape	Analysis of lamina shape	Dornbusch and Andrieu (2010)
	HTPheno	ImageJ plugin for morphological shoot measurements	Hartmann et al. (2011)
	LEAF-GUI	Analysis of leaf vein structure	Price et al. (2011)
	LemnaTec 3D Scanalyzer	Comprehensive platform for analysis of color, shape, size, and architecture	Golzarian et al. (2011), http://www.lemnatec.com/
Seeds/grain	WinSEEDLE	Volume and surface area measurements of seeds and needles	http://www.regent.qc.ca/products/needle/WinSEEDLE.html
	SHAPE	Quantitative evaluation of shape parameters	Iwata and Ukai (2002), Iwata et al. (2010)
	ImageJ	General image analysis software for area, size, and shape; applied to grain	Herridge et al. (2011), http://rsb.info.nih.gov/ij/
	GROWSCREEN-Rhizo	Simultaneous analysis of growth rate, leaf area, and root growth	Nagel et al. (2012)
	SmartGrain	High-throughput measurement of seed shape	Tanabata et al. (2012)

Table 2.2. Selected image analysis software programs available for high throughput phenotyping.

Access, Storage, and Management of Phenotypic Data (This sub-section initially authored by Genevieve DeClerck)

It is clear then, that much of next-generation phenotyping will be done at the intersection of the fields of biology, engineering, and computer science. Progress in developing technology in these disciplines that empowers next-generation phenotyping strategies is moving forward rapidly. However congruent with the progress in the capability to collect high-throughput phenotypic data, is the growing problem of managing these data sets in ways that empower value extraction. Retrofitting a lab to handle the rapid influx of phenotype data could require significant investment in facilities, device control systems and computational resources.

For experiments that are only measuring a few traits on a panel of germplasm, setting up a local (customized) phenotyping system in-house might be practical; but in such cases, a laboratory information management system (LIMS) or local database may be needed to manage the high volumes of phenotypic information. Generally there are no 'off the shelf' solutions that can be applied universally, so some computer expertise will be needed for data management. Even so, organizing that information into a "phenome" is challenging because of the continuous, multi-faceted, and interpretive nature of what a phenotypic observation is, contrasted with the "discrete" nature of genotypic data, which can be abstracted into a single alphabetical character (National Science Foundation 2011).

Beyond the technologies used to run, collect and digest large scale phenotypic evaluations, the field of phenomics faces similar bottlenecks that genomics has been grappling with as the drop in cost of DNA sequencing outpaces the cost of hard drive data storage (Stein 2010). Though there is ample exploration that can be done on genomic data alone, for many plant researchers, associating and enriching genotype data with phenotypic manifestations contextualized by the field environment is a vital part of gaining true biological insight and solving agronomic problems.

The storage of phenotype data at this scale has become a sub-discipline all its own and some projects are dealing with it quite well. There are many public databases that have been working to organize and collate plant phenotype data (Lai et al. 2012; Table 2.3), but most only have the current capacity to present free-text phenotypic descriptions of mutants, e.g. SoyBase (Grant et al. 2010) and MaizeGDB (Schaeffer et al. 2011). Some crop databases have tried to move beyond this paradigm by including functionality for the management of phenotypic measurements, predominantly from either managed field trials or GWA studies (e.g. T3 Triticeae Toolbox (http://triticeaetoolbox.org), Panzea (Canaran et al. 2008), and Gramene's Diversity module (Chen et al. 2010). They are also among a number of projects preparing for an increasing amount of association data emerging from the marriage of powerful genomic information with next-generation phenotyping. One effort is NCBI's dbGAP (Mailman et al. 2007), which was created as a public repository for phenotypes, genotypes and the associations between them. Currently, however, dbGAP only accepts human data.

Taxa	Database	Organism(s)/taxa	Website	References
Plant	MaizeGDB	Maize	http://www.maizegdb.org/	Schaeffer et al. (2011)
	Panzea	Maize; Teosinte	http://www.panzea.org/	Canaran et al. (2008)
	PHENOPSIS DB	*Arabidopsis*	http://bioweb.supagro.inra.fr/phenopsis/	Juliette et al. (2011)
	Gramene Diversity Module	*Arabidopsis*; rice; maize; sorghum; wheat	http://www.gramene.org/db/diversity/diversity_view	Chen et al. (2010)
	IonomicsHub	Arabidopsis; rice; yeast; soybean; maize	http://www.ionomicshub.org/home/PiiMS	Baxter et al. (2007)
	Oryza Tag Line (OTL)	Rice	http://oryzatagline.cirad.fr/	Larmande et al. (2008)
	Rice Mutant database (RMD)	Rice	http://rmd.ncpgr.cn/	Zhang et al. (2006)
	T3 Triticeae Toolbox	Wheat; barley	http://triticeaetoolbox.org/	Blake et al. (2012)
	Tomato Mutant Database	Tomato	http://zamir.sgn.cornell.edu/mutants/	Menda et al. (2004)
	SGN	Solanaceae species	http://solgenomics.net/	Bombarely et al. (2011)
Non-plant	PhenomicDB	Human; mouse; fruit fly; yeast; zebrafish; slimemold; nematode	http://www.phenomicdb.de/	Kahraman et al. (2005)
	MGI	Mouse	http://www.informatics.jax.org/	Eppig et al. (2012)
	The Phenoscape project	Numerous	http://kb.phenoscape.org	Mabee et al. (2012)
	dbGAP	Human	http://www.ncbi.nlm.nih.gov/gap	Mailman et al. (2007)

Table 2.3: A selection of integrated storage and maintenance systems. These systems are designed to enable query-based approaches to characterizing variation in phenotypic data and the relationship it shares with genotypic data.

There are a few database projects that specialize specifically in plant phenomics data, and deserve to be highlighted. The first example of these is PHENOPSIS DB (Juliette et al. 2011), which mainly houses information regarding the growth response of *Arabidopsis thaliana* to various environmental conditions. The database is populated with phenotype information extracted from images and measurements collected automatically in specialized growth chambers. The collaborative international network for ionomics; www.ionomicshub.org; (Baxter et al. 2007) is a second example that hosts ICP-mass spectrometry ionomics data for thousands of *Arabidopsis*, rice, and yeast samples with the goal to facilitate the understanding of response mechanisms in plants to various nutrient availabilities and/or abiotic toxicities.

Additionally, there are other efforts in human and mouse genomics research that could serve as useful models for continued development in the plant phenomics domain. Mouse Genomics Informatics (MGI), (www.informatics.jax.org) is comprised of several database projects, including the Mouse Genome Database (MGD), (Eppig et al. 2012) and houses a variety of tools for searching and browsing large phenotype data sets. PhenomicDB is another, multi-species (primarily human, mouse, fruit fly, and yeast) resource designed to empower "comparative phenomics" (Kahraman et al. 2005). The Nutritional Phenotype Database (Van Ommen et al. 2010) is a third, which focuses on human nutritional phenotype data. The DbNP even goes a step further than most databases by emphasizing the importance of the characterization and unification of experimental designs and allows for finer grained storage and searching of protocol parameters. The DbNP project recently announced that it will further expand the scope

of the resource to include management of environmental plant studies (www.dbnp.org).

One important feature shared among many current databases organizing phenotype data is the use of controlled vocabularies known as 'ontological terms.' Ontologies are sets of defined keywords that can be used as tags to qualify and describe features related to biological data points and data sets. Such ontological terms can be used to describe traits, genes, environments, or taxonomy. As an example, one might use the hierarchy of terms "growth and development=>shoot development=>inflorescence development=>flower development=>flowering time=>days to flower" to describe formally what is colloquially referred to as simply "flowering time." While this is an arguably simple example, it is not difficult to imagine the complexity that ensues when trying to use ontologies to describe more complex molecular pathways. Usage of ontologies is a critical step toward making diverse and rapidly growing collections of biological data searchable, and accessible to computational algorithms. The Open Biological and Biomedical Ontologies Foundry (OBO Foundry) (Smith et al. 2007) has emerged as an important centralized repository for plant and animal ontological collections, with the goal of increasing standardization and maximizing interoperability between ontologies. For plant data the most commonly used ontologies include the Plant Ontology (PO), (Avraham et al. 2008, Jaiswal et al. 2005) the Plant Trait Ontology (TO), Plant Environment Ontology (EO), and the Phenotypic Quality Ontology (PATO).

Unfortunately, all research groups do not universally adopt usage of these community standards and without a critical mass of "buy-in," their benefits cannot be fully realized. Also a great deal of time and resources go in to the curation and maintenance of ontologies and projects rely on term-based grant funding, which is not always reliable.

In order to meet the demand imposed by the upscaling of phenotypic data production, sophisticated computational methods will need to be employed. Phenotype data is complex and highly context sensitive, and crucial information can potentially be lost when data descriptions are flattened down to only a few ontology terms. As a way of dealing with this complexity some groups have been exploring the potential of the Semantic Web (Lee et al. 2001) to expand the dimensionality of stored biological data in order to more effectively mine the enormous volume of descriptive data available in the literature (Vision et al. 2011). The Phenoscape project (kb.phenoscape.org; (Mabee et al. 2012) has been working on developing semantical search algorithms capable of linking biological data by relationships between ontological terms and by similarities found between free-text descriptions. Data are characterized as statements of fact, where there is a subject (e.g. "a floret"), a predicate (e.g. "has the color") and an object (e.g. "white"). Capturing phenotypic metadata using this approach adds some of the necessary dimensionality for unlocking biological meaning using linguistic and intuitive tool sets.

Analysis, Adjustment, and Value extraction of Phenotypic Data (This sub-section initially authored by Dr. Anthony Greenberg, Ph.D)

Independent of the effort involved to both collect and appropriately manage high-throughput phenotype data, the data sets themselves are only as valuable as the analyses that can be performed on them. Great care must be taken to make accurate inferences from the data in order to correctly characterize the genotype-phenotype relationship. Correct estimations of genetic gain from selection, for example, depend heavily on accurate estimates of heritability and the covariance among phenotypes (Dickerson 1955). Because none of these parameters are directly observable, they must be estimated from data using a variety of statistical models.

While the methods for measuring phenotypic data are becoming more sophisticated and the ability to catalog and query data across experimental designs is becoming more achievable, the precision of such data will always be limited by inherent biological noise. This biological noise is unavoidable and is even present under the most controlled experimental conditions. These fluctuations can be local, affecting single organisms, or more general, influencing the whole experiment and modifying the phenotypes for the whole replicate population. Furthermore, where automation is impractical, and a team of researchers is employed to conduct the experiment, individual bias can skew observations, even in cases where subjective criteria are not directly used to measure the phenotypes. These problems are further compounded by the environmental variability that inconsistently affects phenotypic observation over both space and time. Unpredictable environmental conditions can also lead to a fair

amount of missing data, which in turn will limit the statistical power to make inferences about the genotypic contribution to the phenotype. In addition to biological and environmental noise, variable assay quality can introduce further uncertainty and must be accounted for in any statistical models that are used to estimate parameters of interest.

Linear models have long been the mainstay of quantitative genetic experiments, and are the most commonly applied statistical approach to understanding phenotypic variation. Traditionally, these models are fit using a variety of maximum-likelihood approaches (Lynch and Walsh 1997, Sorensen and Gianola 2002). These approaches are popular because they are fast and easy to use, and their long history has resulted in a wide availability of user-friendly software. However, maximum-likelihood methods have a number of serious limitations. Fundamentally, maximum-likelihood model fitting yields point estimates of parameters, ignoring the inherent uncertainty in their values. Parameters are then tested for statistical significance based on a threshold (typically the 5% cut-off) and are excluded from further analysis if they are not "significant." Finally, these statistical tests rely on restrictive assumptions about the distributions of model parameters. These constraints of maximum-likelihood model fitting affect experimental designs and data acquisition procedures, as well as biasing the resulting associations. More pointedly, these estimates perform well only when measurements are extensively replicated and normally distributed. Therefore, a great variety of procedures for data normalization and detection of outliers are necessary in order to meet the assumptions of the model. Unfortunately, these methods are often

poorly motivated from a statistical point of view because they involve arbitrary thresholds for data exclusion. Despite these drawbacks, the speed and prevalence of maximum-likelihood methods make them useful as exploratory data analysis tools even in cases where the resulting estimates are not expected to be robust.

The Bayesian approach to statistical inference is fundamentally different and overcomes many of the limitations imposed by a maximum-likelihood approach. Instead of arriving at single most likely point estimates of parameters, the goal of Bayesian inference is to describe distributions of random variables of interest, taking into account uncertainty in all the other model parameters. This perspective on inference is thus much more in line with biological reality and should be preferable when dealing with phenotype data that have been contextualized by both the genotype and the environment. The drawback of Bayesian inference is its computational complexity. Historically, this complexity has been the disadvantage that held back widespread applications of Bayesian statistics. However, with the dramatic increase in computer power it is now feasible to apply this approach to inference even when the data sets are large and multi-faceted. Furthermore, computer packages that make model building and analysis relatively simple and accessible to researchers without a programming background are starting to make an appearance (Lunn et al. 2009, Plummer 2003). Bayesian formulations of the standard quantitative genetic models have been extensively studied (Sorensen and Gianola 2002), but these models can be computationally inefficient for large data sets. This is true for maximum-likelihood as well, but because Bayesian estimation involves the extra step of estimating full

distributions rather than just point estimates of parameters the computational problems are particularly acute.

Appreciable improvements in computational stability and efficiency can be achieved by re-formulating the standard linear models in a hierarchical framework (Gelman et al. 2003, Gelman and Hill 2007). This framework is popular in the analyses of sociological data, and is now achieving more currency in genetics (Greenberg et al. 2010, Lenarcic et al. 2012). The basic idea is that quantitative-genetic experiments are inherently structured. For example, when an inbred line is evaluated in a number of environments, environmental effects can be nested within genotypic effects. Such nesting improves computational efficiency, increases power by incorporating data-driven pooling of observations from replicates (Gelman et al. 2003, Gelman and Hill 2007, Greenberg et al. 2010), and aids in biological interpretation of the results. Nesting environmental effects within genotypes has the added convenience of allowing the direct modeling of genotype by environment interactions (GxE) simply by estimating the regression slopes as they vary between inbred lines.

In cases where even modest numbers of outlier observations are present, Bayesian hierarchical models also out-perform similar maximum-likelihood approaches (Greenberg et al. 2010). Furthermore, it is straight forward to expand hierarchical models to include non-normal data (Gelman et al. 2003, Gelman and Hill 2007, Greenberg et al. 2010), handle unbalanced designs (Gelman et al. 2003, Greenberg et al. 2010), incorporate variable assay quality (Greenberg et al. 2010, Greenberg et al.

2011), account for outlier observations without using arbitrary thresholds to exclude them from the data (Greenberg et al. 2010, Greenberg et al. 2011, Lenarcic et al. 2012), and interrogate phenotypic networks by extending the analyses to multiple phenotypes through multivariate modeling (Greenberg et al. 2011). Finally, because the Bayesian approach integrates over the inherent uncertainty in a system and borrows power across the experiment through hierarchical modeling, it reduces the need for extensive biological replicates, and therefore maximizes the number of lines that can be evaluated in a given study (Greenberg et al. 2010).

That being said, while Bayesian hierarchical models are robust to many problems in experimental design and data acquisition, it is still a good idea to follow best practices when embarking on a quantitative genetic experiment. Certain problems, such as putting all replicates for a line in a single block, lead to complete confounding of variables that cannot be resolved by any statistical treatment. Although it is possible to incorporate non-Gaussian data into Bayesian models, these extensions are typically computationally more expensive. For example, when modeling categorical data one attempts to estimate an underlying continuous distribution that would yield the observed data when coerced to being discrete. Converting quantitative phenotypes (for example, the fraction of a plant tissue affected by disease) to an index (susceptibility class) leads to loss of information and an increase in model complexity. Likewise, summarizing replicate observations and reporting only means can lead to either increased noise when outliers are present or unwarranted precision. Such short cuts were defensible in the past, when computational power and storage capacity to handle

large data sets was limited, but this is no longer the case and the data should be reported as “raw" as possible, and then modeled explicitly.

Germplasm Development and Distributed Phenotyping Networks

Advances in phenotyping and genotyping technology, as well as data storage, and computational capacity are opening many new opportunities to extract meaningful inferences from even noisy biological data. New statistical models that account for biological uncertainty and estimate values of direct interest, rather than those dictated by computational convenience, promise to aid in the achievement of this goal. However, the value of any progress that may be gained through the marriage of next-generation phenotyping with modern genomic tools is predicated on the availability of diverse germplasm and genetically well-defined populations. Indeed associating genotype with phenotype in ways that address hypothesis-driven questions and empower crop improvement depends on the availability of appropriate germplasm resources to address specific questions.

The preservation of plant biodiversity in publicly available, international germplasm collections is of central importance to our quest to understand natural variation and to utilize that variation to meet the future needs of the planet. It is not unimaginable that we will be able to genomically characterize most of the accessions in the world’s repositories of genetic resources, but the sheer size of these collections, the broad range of adaptation they represent, import-export restrictions, and the genetic redundancy housed within their ranks presents a challenge for direct phenotypic

evaluation. Targeted subsets of this variation need to be assembled so that available phenotyping resources can be efficiently used to evaluate them, taking advantage of economies of scale wherever possible (Glaszmann et al., 2010; McCouch et al., 2012). The development of shared populations with publically available, high-resolution genotype data will be critical for permitting the kind of distributed phenotyping necessary to understand genotype-phenotype relationships (Valdar et al. 2006). Examples of research communities that have developed these kinds of publicly shared germplasm resources include rice (Zhao et al. 2011a), maize (Yu et al. 2008), wheat (Neumann et al. 2011) arabidopsis (Atwell et al. 2010), sorghum (Mitchell et al. 2008), barley (Pasam et al. 2012) and many other species (Zhu et al. 2008). The availability of these resources makes it possible for multiple researchers to interrogate the same genetic materials, phenotyping in environments and with technology and analytical expertise that are uniquely available to different research groups. Integrating such vast phenotypic datasets on common germplasm resources in well-structured databases will permit high-end analysis not just of the phenotypes themselves, but also of complex correlated phenotypic networks that represent a more accurate depiction of biological reality.

Additionally, more genetically structured resources such as chromosome segment substitution lines (Ali et al., 2011; Lu et al. 2011, Wang et al. 2012, Fukuoka et al., 2010; Xu et al., 2010; Zhang et al. 2011), MAGIC (multi-parent advanced generation inter-cross) populations (Huang et al. 2012, Huang et al. 2011, Rakshit et al. 2012), and nested association mapping (NAM) populations (Yu et al. 2008) will permit the

interrogation of natural variation in elite genetic backgrounds that may be intractable otherwise. These genetically structured populations partition the variation in ways that facilitate the identification of exotic alleles that may have a significant impact on a phenotype of interest, but only when introgressed into the elite background. They also expedite the subsequent use of these resources as parents in a breeding program, helping expand the range of genetic variation available in an elite gene pool and opening up new opportunities to utilize natural variation to drive crop improvement.

Conclusions

Ever since the first published QTL analysis (Sax 1923) genetics as a discipline has endeavored to shed light on the complexities of phenotypic variation. For most of recent memory, progress in understanding the genetic architecture of complex traits has been driven by advances in genotyping technology. As a clear picture of the genome emerges, a renewed focus on understanding the nature of phenotypes will be necessary for continued progress.

We have discussed the role of phenotyping in gene-discovery and crop improvement through both GWAS and genomic selection, and we have attempted to understand the complexities incumbent on the association of genotype with phenotype under diverse and variable environmental conditions. We considered strategies that permit the collection of phenotypic data in quantitative ways as well as the development of modular technologies to accommodate the changing needs and opportunities of phenotyping in the future. We have pondered best practices for storing, cataloging,

managing, and disseminating this information within a community, and suggested how this data might be combined with cutting edge statistical analysis to leverage increased computing capacity (Figure 2.2). To conclude, we consider where some of the current deficiencies lie and highlight a few questions that still need answers.

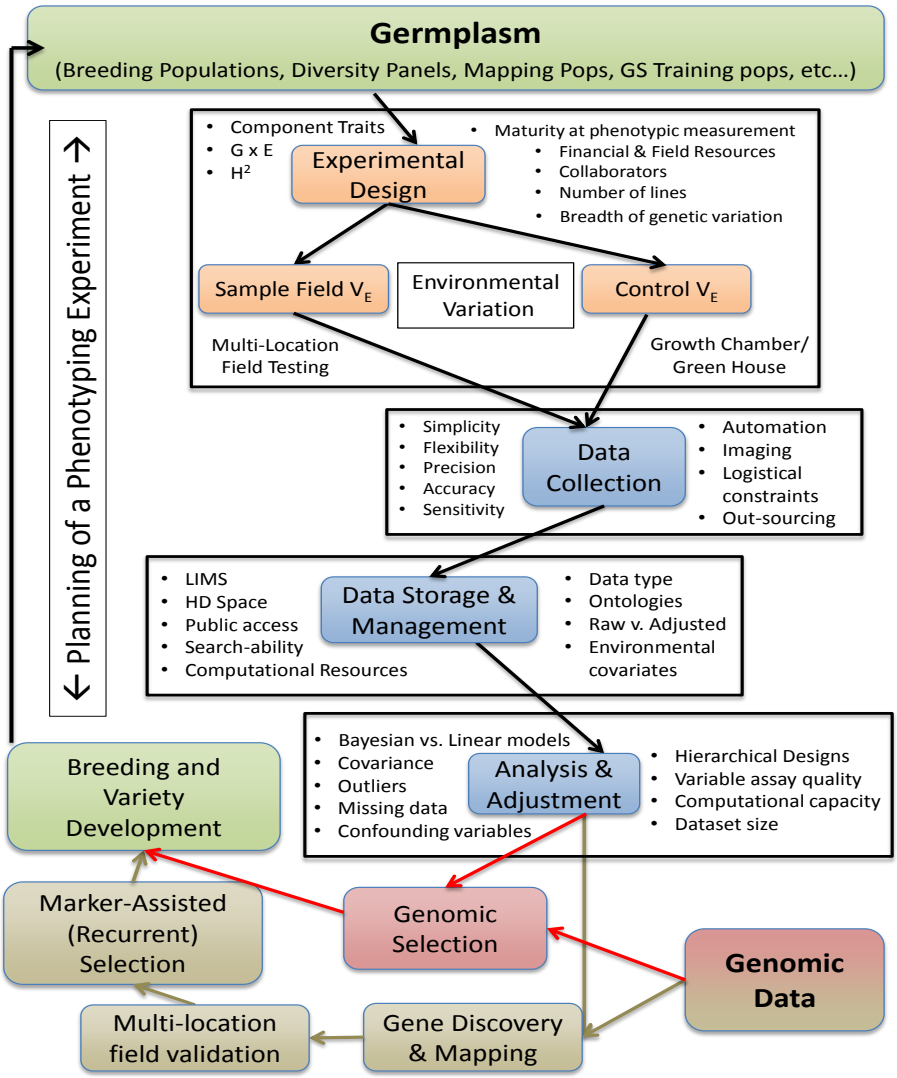


Figure 2.2. Outline of decision making considerations for genetic experimentation. When combined with high-throughput genotyping on shared germplasm resources, and done in geographically distributed collaborative networks, next-generation phenotyping can empower both gene discovery and crop improvement. Central to that capacity is the careful and judicious use of modular technologies and managed environments. The use of standardized ontologies and Bayesian analysis then create a controlled vocabulary for describing the data and provide a way to integrate results across experiments by accounting for the unique signatures of biological noise generated by environmental covariates.

Genotyping, while closing in on understanding the full extent of allelic variation in major crop species, is still years away from delivering on the quest to catalogue the world's collection of DNA variants for an entire species. This requires assembly of multiple *de novo* reference genomes and re-sequencing of thousands of diverse lines to identify all of the SNPs, copy number variants, and other forms of DNA and epigenetic variation within a gene pool. As that information is generated, researchers will seek to annotate the functional significance of the SNPs and insertion/deletion polymorphisms, and design databases that can host this information and make it accessible and query-able for the research community. This is a real challenge because many functional variants do not fall within gene models, but are found as inter-genic regulatory elements or may condition gene expression through epigenetic pathways that contribute to quantitative phenotypic variation (Ding et al. 2012, Loehlin et al. 2010, Salvi et al. 2007, Zhou et al. 2012, Zhu and Deng 2012). This challenge also highlights the value of positional cloning to verify the functional nucleotide polymorphisms (FNP) rather than taking a candidate gene approach, as the FNP may not be found within a gene model at all. Additionally, for many years to come, the identification and characterization of rare alleles will remain a priority, despite the fact that both GWAS and genomic selection have little power to detect their contributions to phenotypic variation.

Algorithms for optimizing signal to noise ratios in phenotypic experiments, pipelines for identifying GWAS peaks and extracting meaningful lists of candidate genes underlying those peaks are needed to help standardize association mapping studies.

Functional annotation of QTL alleles and correspondence to the germplasm samples in which they are found would help link genetics research with breeding applications. Better tools for SNP haplotype visualization and management of high-volume SNP data need to be integrated into software platforms to facilitate the identification of functionally relevant SNPs that can be used for marker assisted selection and as fixed variables in genomic prediction. As more and more phenotype data are collected and databased, tools to facilitate our understanding of intersecting phenotypic networks will shed light on the complex relationships within and between phenotypes (Yin and Struik 2008). This information will provide important insights about selection trade-offs and phenotypic correlations that are relevant to variety development and plant breeding.

Major questions about phenotypic variation, which we currently have limited capacity to answer, include: How does variation in regulatory elements manifest itself in the phenotype? Which environmental variables act as signals that regulate these genes and how do different allelic variants recognize those signals? What is the role of epistasis and epigenetics in determining phenotypic variation, or in phenotypic plasticity?

Approaching many of those questions will require more refined strategies of collecting and managing phenotype data. Many of the considerations that need to be addressed before making decisions about defining a phenotyping approach include: How easy is it to evaluate the phenotype? How quantitative is that measurement? Can the process be automated? If so, does it make economic sense to do so? What value would

automation bring? What indirect factors will influence the phenotypic measurement? Can they be quantified? How much storage capacity do I need to maintain the raw or processed phenotypic data? How will the data be organized so that it is both query-able and understandable? What data processing needs must be considered before the phenotype is biologically meaningful? Do I have the skills in-house or appropriate collaborators in place to realize a sophisticated analysis of the data? Answers to these questions will depend entirely on the purpose and intention of collecting phenotypic data to start with, and of course the nature of the phenotype itself.

The phenotype of an organism is fundamentally a manifestation of a genotype's interaction with the environment. With increased allocation of funding and intellectual investment over the next decade, advances in phenotyping will enhance our ability to associate that data with the genotypic and environmental variables to simultaneously and synergistically drive gene discovery efforts aimed at understanding the genetic basis of quantitative phenotypic variation and fuel the development of genomic prediction models for crop improvement. As these two drivers of genetic analysis feed into each other, not only will tremendous gains be made in comprehending the biology of plants, but also we will ensure continued advancement in crop improvement aimed at meeting the demands of a growing population.

Acknowledgments

I would like to acknowledge Lukas Mueller of the Boyce Thompson Institute for Plant Research and SOL Genomics Network (SOL; solgenomics.org), Dave

Matthews of USDA-PWA and the Triticeae Toolbox (triticeaetoolbox.org), Jean-Luc Jannink of the USDA-ARS for valuable discussion and insight, Michael Gore of the USDA-ARS for helpful review of the manuscript, and Cheryl Utter for help with formatting.

REFERENCES

Ali, M., Sanchez, P.L., Yu, S., Lorieux, M., Eizenga, G.C. 2010. Chromosome segment substitution lines: A powerful tool for the introgression of valuable genes from wild species of rice (*Oryza* spp.). Rice. 3:218–234.

Armengaud P, Zambaux K, Hills A, Sulpice R, Pattison RJ, Blatt MR, Amtmann A (2009) EZ‐Rhizo: integrated software for the fast and accurate measurement of root system architecture. Plant J 57:945-956

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465:627-631

Avraham S, Tung CW, Ilic K, Jaiswal P, Kellogg EA, McCouch S, Pujar A, Reiser L, Rhee SY, Sachs MM (2008) The plant ontology database: A community resource for plant structure and developmental stages controlled vocabulary and annotations. Nucleic Acids Res 36:D449-D454

Bachlava E, Taylor CA, Tang S, Bowers JE, Mandel JR, Burke JM, Knapp SJ (2012) SNP discovery and development of a high-density genotyping array for sunflower. PLoS One doi:10.1371/journal.pone.0029814

Backhaus A, Kuwabara A, Bauch M, Monk N, Sanguinetti G, Fleming A (2010) LEAFPROCESSOR: a new leaf phenotyping tool using contour bending energy and shape cluster analysis. New Phytol 187:251-261

Bakker M, Manter D, Sheflin A, Weir T, Vivanco J (2012) Harnessing the rhizosphere microbiome through plant breeding and agricultural management. Plant Soil 360:1-13

Bartlett MK, Scoffoni C, Ardy R, Zhang Y, Sun S, Cao K, Sack L (2012a) Rapid determination of comparative drought tolerance traits: using an osmometer to predict turgor loss point. Methods Ecol Evol 3:880-888

Bartlett MK, Scoffoni C, Sack L (2012b) The determinants of leaf turgor loss point and prediction of drought tolerance of species and biomes: a global meta‐analysis. Ecol Lett 15:393-405

Basu P, Pal A, Lynch JP, Brown KM (2007) A novel image-analysis technique for kinematic study of growth and curvature. Plant Physiol 145:305-316

Baxter I, Ouzzani M, Orcun S, Kennedy B, Jandhyala SS, Salt DE (2007) Purdue ionomics information management system. An integrated functional genomics platform. Plant Physiol 143:600-611

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Paterson AH (ed) Molecular dissection of complex traits. CRC Press, New York, New York, pp 145-162

Bennett D, Reynolds M, Mullan D, Izanloo A, Kuchel H, Langridge P, Schnurbusch T (2012) Detection of two major grain yield QTL in bread wheat (Triticum aestivum L.) under heat, drought and high yield potential environments. Theor Appl Genet 125:1473-1485

Berger B, Parent B, Tester M (2010) High-throughput shoot imaging to study drought responses. J Exp Bot 61:3519-3528

Bernier J, Kumar A, Ramaiah V, Spaner D, Atlin G (2007) A Large-Effect QTL for Grain Yield under Reproductive-Stage Drought Stress in Upland Rice. Crop Sci 47:507-518

Bhattacharyya MK (2010) Map-based Cloning of Genes and QTL in Soybean. In: Bilyeu K, Ratnaparkhe MB, Kole C (eds) Genetics, Genomics, and Breeding of Soybean. Science Publishers, Enfield, NY, pp 169-186

Blake VC, Kling JG, Hayes PM, Jannink JL, Jillella SR, Lee J, Matthews DE, Chao S, Close TJ, Muehlbauer GJ (2012) The Hordeum Toolbox: The Barley Coordinated Agricultural Project Genotype and Phenotype Resource. Plant Gen 5:81-91

Blum A (2009) Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. Field Crops Res 112:119-123

Bombarely A, Menda N, Tecle IY, Buels RM, Strickler S, Fischer-York T, Pujar A, Leto J, Gosselin J, Mueller LA (2011) The Sol Genomics Network (solgenomics. net): growing tomatoes using Perl. Nucleic Acids Res 39:D1149-D1155

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Oropeza Rosas M, Rocheford TR, Romay MC, Romero S, Salvo S, Sanchez Villeda H, da Silva HS, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen

MD (2009) The genetic architecture of maize flowering time. Science 325 (5941), 714-718. [DOI:10.1126/science.1174276]

Bylesjö M, Segura V, Soolanayakanahally RY, Rae AM, Trygg J, Gustafsson P, Jansson S, Street NR (2008) LAMINA: a tool for rapid quantification of leaf size and shape parameters. BMC Plant Biol 8:82

Cabrera-Bosquet L, Crossa J, von Zitzewitz J, Dolors Serret M, Araus JL (2012) High-throughput phenotyping and genomic selection: The frontiers of crop breeding converge. J Intergr Plant Bio 54:312-320

Cai X, Molden D, Mainuddin M, Sharma B, Ahmad M, Karimi P (2011) Producing more food with less water in a changing world: Assessment of water productivity in 10 major river basins. Water Int 36:42-62

Calus MPL, Meuwissen THE, Roos APW, Veerkamp RF (2008) Accuracy of genomic selection using different methods to define haplotypes. Genetics 178: 553–561

Campos H, Cooper M, Habben J, Edmeades G, Schussler J (2004) Improving drought tolerance in maize: a view from industry. Field Crops Res 90:19-34

Canaran P, Buckler ES, Glaubitz JC, Stein L, Sun Q, Zhao W, Ware D (2008) Panzea: An update on new content and features. Nucleic Acids Res 36:D1041-D1043

Cavanagh C, Morell M, Mackay I, Powell W (2008) From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Curr Opin Plant Biol 11:215-221

Chen C, DeClerck G, Casstevens T, Youens-Clark K, Zhang J, Ware D, Jaiswal P, McCouch S, Buckler E (2010) The gramene genetic diversity module: A resource for genotype-phenotype association analysis in grass species. Nature Precedings hdl:10101/npre.2010.4645.1

Chen H, He H, Zou Y, Chen W, Yu R, Liu X, Yang Y, Gao YM, Xu JL, Fan LM, Li Y, Li ZK, Deng XW (2011) Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.). Theor Appl Genet doi: 10.1007/s00122-011-1633-5

Clark RT, Famoso AN, Zhao K, Shaff JE, Craft EJ, Bustamante CD, McCouch SR, Aneshansley DJ, Kochian LV (2012) High-throughput two dimensional root system phenotyping platform facilitates genetic analysis of root growth and development. Plant Cell Environ doi: 10.1111/j.1365-3040.2012.02587.x

Clark RT, MacCurdy RB, Jung JK, Shaff JE, McCouch SR, Aneshansley DJ, Kochian LV (2011) Three-dimensional root phenotyping with a novel imaging and software platform. Plant Physiol 156:455-465

Clarke JM, DePauw RM, Townley-Smith TF (1992) Evaluation of methods for quantification of drought tolerance in wheat. Crop Sci 32:723-728

Collard BCY, Mackill DJ (2008) Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. Phil Trans R Soc B 363:557-572

Crouch J, Payne T, Dreisigacker S, Wu H, Braun H (2009) Improved discovery and utilization of new traits for breeding. In: Dixon JM (ed) Wheat Facts and Futures 2009. CIMMYT, pp 42-51

Czyczyło-Mysza I, Marcińska I, Skrzypek E, Chrupek M, Grzesiak S, Hura T, Stojałowski S, Myśków B, Milczarski P, Quarrie S (2011) Mapping QTLs for yield components and chlorophyll a fluorescence parameters in wheat under three levels of water availability. Plant Gen Res 9:291-295

De Boever J, De Brabander D, De Smet A, Vanacker J, Boucqué CV (1993) Evaluation of physical structure. 2. Maize silage. J Dairy Sci 76:1624-1634

Dickerson G (1955) Genetic slippage in response to selection for multiple objectives. Cold Spring Harb Symp Quant Biol 20:213-224

Ding J, Lu Q, Ouyang Y, Mao H, Zhang P, Yao J, Xu C, Li X, Xiao J, Zhang Q (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. Proc Natl Acad Sci 109:2654-2659

Dodig D, Zoric M, Kobiljski B, Savic J, Kandic V, Quarrie S, Barnes J (2012) Genetic and Association Mapping Study of Wheat Agronomic Traits Under Contrasting Water Regimes. Int J Mol Sci 13:6167-6188

Dornbusch T, Andrieu B (2010) Lamina2Shape—An image processing tool for an explicit description of lamina shape tested on winter wheat (*Triticum aestivum*L.). Comput Electron Agric 70:217-224

Elshire RJ, Glaubitz JC, Sun Q, Poland Ja, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one 6:e19379

Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE (2012) The Mouse Genome Database (MGD): Comprehensive resource for genetics and genomics of the laboratory mouse. Nucleic Acids Res 40:D881-D886

Famoso AN, Clark RT, Shaff JE, Craft E, McCouch SR, Kochian LV (2010) Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal aluminum tolerance and investigations into rice aluminum tolerance mechanisms. Plant Physiol 153:1678-1691

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112:1164-1171

Ferguson ME, Hearne SJ, Close TJ, Wanamaker S, Moskal WA, Town CD, de Young J, Marri PR, Rabbi IY, de Villiers EP (2011) Identification, validation and high-throughput genotyping of transcribed gene SNPs in cassava. Theor Appl Genet. doi: 10.1007/s00122-011-1739-9

Fernie AR, Schauer N (2009) Metabolomics-assisted breeding: A viable option for crop improvement? Trends Genet 25:39-48

Foolad MR, Panthee DR (2012) Marker-assisted selection in tomato breeding. Crit Rev Plant Sci 31:93-123

French A, Ubeda-Tomás S, Holman TJ, Bennett MJ, Pridmore T (2009) High-throughput quantification of root growth using a novel image-analysis tool. Plant Physiol 150:1784-1795

Frison E, Cherfas J, Hodgkin T (2011) Agricultural biodiversity is essential for a sustainable improvement in food and nutrition security. Sustainability 3:238-253

Fukuoka S, Nonoue Y, Yano M (2010) Germplasm enhancement by developing advanced plant materials from diverse rice accessions. Breeding Science 60:509-517. DOI: 10.1270/jsbbs.60.509.

Galkovskyi T, Mileyko Y, Bucksch A, Moore B, Symonova O, Price CA, Topp CN, Iyer-Pascuzzi AS, Zurek PR, Fang S (2012) GiA Roots: Software for the high throughput analysis of plant root system architecture. BMC Plant Biol 12:116

Gelman A, Carlin JB, Stern HS, Rubin DB (2003) Bayesian data analysis. Chapman & Hall CRC Press

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York

Ghimire KH, Quiatchon LA, Vikram P, Swamy B, Dixit S, Ahmed H, Hernandez JE, Borromeo TH, Kumar A (2012) Identification and mapping of a QTL (*qDTY*1.1) with a consistent effect on grain yield under drought. Field Crops Res 131:88-96

Gilbert ME, Zwieniecki MA, Holbrook NM (2011) Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. J Exp Bot 62:2875-2887

Glaszmann JC, Kilian B, Upadhyaya HD, Varshney RK (2010) Accessing genetic diversity for crop improvement. Current Opinion in Plant Biology doi: 10.1016/j.pbi.2010.01.004

Golabadi M, Arzani A, Mirmohammadi Maibody SAM, Sayed Tabatabaei B, Mohammadi S (2011) Identification of microsatellite markers linked with yield components under drought stress at terminal growth stages in durum wheat. Euphytica 177:207-221

Goltsev V, Zaharieva I, Chernev P, Kouzmanova M, Kalaji HM, Yordanov I, Krasteva V, Alexandrov V, Stefanov D, Allakhverdiev SI (2012) Drought-induced modifications of photosynthetic electron transport in intact leaves: Analysis and use of neural networks as a tool for a rapid non-invasive estimation. Biochim Biophys Acta 1817:1490-1498

Golzarian MR, Frick RA, Rajendran K, Berger B, Roy S, Tester M, Lun DS (2011) Accurate inference of shoot biomass from high-throughput images of cereal plants. Plant Methods 7:2

Gordon D, Finch SJ (2005) Factors affecting statistical power in the detection of genetic association. J Clin Invest 115:1408-1418

Granier C, Aguirrezabal L, Chenu K, Cookson SJ, Dauzat M, Hamard P, Thioux JJ, Rolland G, Bouchier‐Combaud S, Lebaudy A (2006) PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. New Phytol 169:623-635

Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res 38:D843-D846

Greenberg AJ, Hackett SR, Harshman LG, Clark AG (2010) A hierarchical Bayesian model for a novel sparse partial Diallel crossing design. Genetics 185:361-373

Greenberg AJ, Hackett SR, Harshman LG, Clark AG (2011) Environmental and genetic perturbations reveal different networks of metabolic regulation. Mol Syst Biol 7:563

Guo Z, Tucker DM, Lu J, Kishore V, Gay G (2012) Evaluation of genome-wide selection efficiency in maize nested association mapping populations. Theor Appl Genet 124:261-275

Hartmann A, Czauderna T, Hoffmann R, Stein N, Schreiber F (2011) HTPheno: An image analysis pipeline for high-throughput plant phenotyping. BMC Bioinformatics 12:148

Hayes B, Bowman P, Chamberlain A, Goddard M (2009) Invited review: Genomic selection in dairy cattle: Progress and challenges. J Dairy Sci 92:433-443

Heffner EL, Jannink J, Sorrells ME (2010) Genomic selection accuracy using multifamily prediction models in a wheat breeding program. The Plant Genome 4:65-75

Heffner EL, Sorrells ME, Jannink J (2009) Genomic selection for crop improvement. Crop Sci 49:1-12

Herridge RP, Day RC, Baldwin S, Macknight RC (2011) Rapid analysis of seed size in Arabidopsis for mutant and QTL discovery. Plant Methods 7:3

Heslot N, Sorrells ME, Jannink JL, Yang HP (2012) Genomic selection in plant breeding: A comparison of models. Crop Sci 52:146-160

Houle D, Govindaraju DR, Omholt S (2010) Phenomics: The next challenge. Nat Rev Genet 11:855-866

Huang BE, George AW, Forrest KL, Kilian A, Hayden MJ, Morell MK, Cavanagh CR (2012) A multiparent advanced generation inter-cross population for genetic analysis in wheat. Plant Biotechnol J 10:826-839

Huang X, Paulo MJ, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk FA (2011) Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. Proc Natl Acad Sci 108:4488-4493

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang Q, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961-967

Hyten D, Song Q, Fickus E, Quigley C, Lim J, Choi I, Hwang E, Pastor-Corrales M, Cregan P (2010) High-throughput SNP discovery and assay development in common bean. BMC Genomics. doi: 10.1186/1471-2164-11-475

Ibrahim S, Schubert A, Pillen K, Léon J (2012) QTL analysis of drought tolerance for seedling root morphological traits in an advanced backcross population of spring wheat. International Journal of AgriScience 2:619-629

International Rice Research Institute (1996) Standard Evaluation System for Rice

Iwata H, Ebana K, Uga Y, Hayashi T, Jannink JL (2010) Genome-wide association study of grain shape variation among Oryza sativa L. germplasms based on elliptic Fourier analysis. Mol Breed 25:203-215

Iwata H, Ukai Y (2002) SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. J Hered 93:384-385

Jaiswal P, Avraham S, Ilic K, Kellogg EA, McCouch S, Pujar A, Reiser L, Rhee SY, Sachs MM, Schaeffer M (2005) Plant Ontology (PO): A controlled vocabulary of plant structures and growth stages. Comp Funct Genomics 6:388-397

Jansen M, Gilmer F, Biskup B, Nagel KA, Rascher U, Fischbach A, Briem S, Dreissen G, Tittmann S, Braun S (2009) Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in Arabidopsis thaliana and other rosette plants. Funct Plant Biol 36:902-914

Jia Y, Jannink JL (2012) Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. Genetics 192:1513-1522

Jin L, Lu Y, Shao Y, Zhang G, Xiao P, Shen S, Corke H, Bao J (2010) Molecular marker assisted selection for improvement of the eating, cooking and sensory quality of rice (*Oryza sativa* L.). J Cereal Sci 51:159-164

Juliette F, Myriam D, Vincent N, Nathalie W, Emilie G, Pascal N, Sébastien T, Catherine M, Irène H, Christine G (2011) PHENOPSIS DB: An information system for *Arabidopsis thaliana* phenotypic data in an environmental context. BMC Plant Biology 11:77

Kahraman A, Avramov A, Nashev LG, Popov D, Ternes R, Pohlenz HD, Weiss B (2005) PhenomicDB: A multi-species genotype/phenotype database for comparative phenomics. Bioinformatics 21:418-420

Kholová J, Hash CT, Kakkera A, Kočová M, Vadez V (2010) Constitutive water-conserving mechanisms are correlated with the terminal drought tolerance of pearl millet [Pennisetum glaucum (L.) R. Br.]. J Exp Bot 61:369-377

Kloth KJ, Thoen MPM, Bouwmeester HJ, Jongsma MA, Dicke M (2012) Association mapping of plant resistance to insects. In: Feuillet C, Muehlbauer GJ (eds) Genetics and Genomics of the Triticeae. Elsevier, pp 337-357

Krattinger S, Wicker T, Keller B (2009) Map-based cloning of genes in Triticeae (wheat and barley). Genetics and Genomics of the Triticeae:337-357

Kuhn C, Smith T (1977) Effectiveness of a disease index system in evaluating corn for resistance to maize dwarf mosaic virus. Phytopathology 67:288-291

Kumar S, Bink MCAM, Volz RK, Bus VGM, Chagné D (2012) Towards genomic selection in apple (Malus× domestica Borkh.) breeding programmes: Prospects, challenges and strategies. Tree Genet Genomes 8:1-14

Lai K, Lorenc MT, Edwards D (2012) Genomic databases for crop improvement. Agronomy 2:62-73

Landi P, Giuliani S, Salvi S, Ferri M, Tuberosa R, Sanguineti MC (2010) Characterization of root-yield-1.06, a major constitutive QTL for root and agronomic traits in maize across water regimes. J Exp Bot 61:3553-3562

Larmande P, Gay C, Lorieux M, Périn C, Bouniol M, Droc G, Sallaud C, Perez P, Barnola I, Biderre-Petit C (2008) Oryza Tag Line, a phenotypic mutant database for the Genoplante rice insertion line library. Nucleic Acids Res 36:D1022-D1027

Le Bot J, Serra V, Fabre J, Draye X, Adamowicz S, Pagès L (2010) DART: a software to analyse root system architecture and development from captured images. Plant Soil 326:261-273

Lee TB, Hendler J, Lassila O (2001) The semantic web. Sci Am 284:34-43

Lenarcic AB, Svenson KL, Churchill GA, Valdar W (2012) A general Bayesian approach to analyzing diallel crosses of inbred strains. Genetics 190:413-435

Li Q, Yang X, Bai G, Warburton ML, Mahuku G, Gore M, Dai J, Li J, Yan J (2010) Cloning and characterization of a putative GS3 ortholog involved in maize kernel development. Theor Appl Genet 120:753-763

Liu S, Pumphrey MO, Gill BS, Trick HN, Zhang JX, Dolezel J, Chalhoub B, Anderson JA (2008) Toward positional cloning of Fhb1, a major QTL for Fusarium head blight resistance in wheat. Cereal Res Commun 36:195-201

Liu Y, Subhash C, Yan J, Song C, Zhao J, Li J (2011) Maize leaf temperature responses to drought: Thermal imaging and quantitative trait loci (QTL) mapping. Environ Exp Bot 71:158-165

Lobet G, Pagès L, Draye X (2011) A novel image-analysis toolbox enabling quantitative analysis of root system architecture. Plant Physiol 157:29-39

Loehlin DW, Oliveira DCSG, Edwards R, Giebel JD, Clark ME, Cattani MV, van de Zande L, Verhulst EC, Beukeboom LW, Muñoz-Torres M (2010) Non-coding changes cause sex-specific wing size differences between closely related species of Nasonia. PLoS Genet 6:e1000821

Lopes MS, Araus JL, Van Heerden PDR, Foyer CH (2011) Enhancing drought tolerance in C4 crops. J Exp Bot 62:3135-3153

Lu MY, Li XH, Shang AL, Wang YM, Xi ZY (2011) Characterization of a set of chromosome single-segment substitution lines derived from two sequenced elite maize inbred lines. Maydica 56:399-407

Lunn D, Spiegelhalter D, Thomas A, Best N (2009) The BUGS project: Evolution, critique and future directions. Stat Med 28:3049-3067

Lynch M, Walsh B (1997) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA

Mabee P, Balhoff J, Dahdul W, Lapp H, Midford P, Vision T, Westerfield M (2012) 500,000 fish phenotypes: The new informatics landscape for evolutionary and developmental biology of the vertebrate skeleton. J Appl Ichthyol 28:300-305

Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L (2007) The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 39:1181-1186

Manolio TA and Collins FS (2009) The HapMap and Genome-Wide Association Studies in Diagnosis and Therapy. Annu. Rev. Med. 2009. 60:443–56

Masuka B, Araus JL, Das B, Sonder K, Cairns JE (2012) Phenotyping for Abiotic Stress Tolerance in Maize. J Integr Plant Biol 54:238-249

Maughan P, Smith S, Fairbanks D, Jellen E (2011) Development, characterization, and linkage mapping of single nucleotide polymorphisms in the grain Amaranthus (*Amaranthus* sp.). The Plant Genome 4:92-101

McCouch SR, McNally KL, Wang W, Sackville Hamilton R (2012) Genomics of gene banks: A case study in rice. Am J Bot 99:407-423

Menda N, Semel Y, Peled D, Eshed Y, Zamir D (2004) In silico screening of a saturated mutation library of tomato. Plant J 38:861-872

Messmer R, Fracheboud Y, Bänziger M, Vargas M, Stamp P, Ribaut JM (2009) Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. Theor Appl Genet 119:913-930

Mir RR, Zaman-Allah M, Sreenivasulu N, Trethowan R, Varshney RK (2012) Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. Theor Appl Genet 125:625-645

Mitchell SE, Casa AM, Tuinstra MR, Brown PJ, Pressoir G, Rooney WL, Franks CD, Kresovich S (2008) Community resources and strategies for association mapping in sorghum. Crop Sci 48:30-40

Molina-Cano J (1987) The EEC Barley and Malt Committee index for the evaluation of malting quality in barley and its use in breeding. Plant breeding 98:249-256

Montes JM, Melchinger AE, Reif JC (2007) Novel throughput phenotyping platforms in plant genetic studies. Trends Plant Sci 12:433-436

Munoz P, Resende M, Peter G, Huber D, Kirst M, Quesada T (2011) Effect of BLUP prediction on genomic selection: practical considerations to achieve greater accuracy in genomic selection. BMC Proceedings 5:49

Naeem A, French AP, Wells DM, Pridmore TP (2011) High-throughput feature counting and measurement of roots. Bioinformatics 27:1337-1338

Nagel KA, Putz A, Gilmer F, Heinz K, Fischbach A, Pfeifer J, Faget M, Blossfeld S, Ernst M, Dimaki C (2012) GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. Funct Plant Biol 39(11):891-904

Nakaya A, Isobe SN (2012) Will genomic selection be a practical method for plant breeding?. Ann Bot 110:i-iv-doi:10.1093/aob/mcs229

National Science Foundation (2011) Phenomics: Genotype to phenotype, a report of the NIFA-NSF phenomics workshop.

Neumann K (2013) Using Automated High-Throughput Phenotyping using the LemnaTec Imaging Platform to Visualize and Quantify Stress Influence in Barley. PAG XXI Jan 11-16, 2013:San Diego, CA

Neumann K, Kobiljski B, Denčić S, Varshney R, Börner A (2011) Genome-wide association mapping: a case study in bread wheat (Triticum aestivum L.). Mol Breed 27:37-58

Ogburn R, Edwards EJ (2012) Quantifying succulence: a rapid, physiologically meaningful metric of plant water storage. Plant, Cell Environ 35:1533-1542

Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B, Graner A (2012) Genome-wide association studies for agronomical traits in a world wide spring barley collection. BMC Plant Biology 12:16. DOI: 10.1186/1471-2229-12-16

Paux E, Faure S, Choulet F, Roger D, Gauthier V, Martinant JP, Sourdille P, Balfourier F, Le Paslier MC, Chauveau A (2010) Insertion site-based polymorphism markers open new perspectives for genome saturation and marker ‐ assisted selection in wheat. Plant Biotechnol J 8:196-210

Pereyra-Irujo GA, Gasco ED, Peirone LS, Aguirrezábal LAN (2012) GlyPh: a low-cost platform for phenotyping plant growth and water use. Funct Plant Biol 39:905-913

Perlin MW, Lancia G, Ng SK (1995) Toward fully automated genotyping: genotyping microsatellite markers by deconvolution. Am J Hum Genet 57:1199

Pieruschka R, Poorter H (2012) Phenotyping plants: genes, phenes and machines. Funct Plant Biol 39:813-820

Plummer M (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International workshop on distributed statistical computing. http://citeseer.ist.psu.edu/plummer03jags.html

Poland J, Nelson R (2010) In the eye of the beholder: The effect of rater variability and different rating scales on QTL mapping. Phytopathology. doi: 10.1094/PHYTO-03-10-0087

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nature Genet 38(8):904-909. doi:10.1038/ng1847

Price CA, Symonova O, Mileyko Y, Hilley T, Weitz JS (2011) Leaf extraction and analysis framework graphical user interface: segmenting and analyzing the structure of leaf veins and areoles. Plant Physiol 155:236-245

Pritchard JK and Cox NJ (2002) The allelic architecture of human disease genes: common disease – common variant… or not? Human Molecular Genetics 11(20):2427-2423.

Pypers P, Sanginga JM, Kasareka B, Walangululu M, Vanlauwe B (2011) Increased productivity through integrated soil fertility management in cassava-legume

intercropping systems in the highlands Sud-Kivu DR Congo. Field Crops Res 120:76-85

Rafalski JA (2010) Association genetics in crop improvement. Curr Opin Plant Biol 13:174-180

Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet 9:275-280

Rakshit S, Rakshit A, Patil J (2012) Multiparent intercross populations in analysis of quantitative traits. J Genet 91:111-117

Rehman A, Malhotra R, Bett K, Tar'an B, Bueckert R, Warkentin T (2011) Mapping QTL Associated with Traits Affecting Grain Yield in Chickpea (L.) under Terminal Drought Stress. Crop Sci 51:450-463

Reich DE and Lander ES (2001) On the allelic spectrum of human disease. Trends in Genetics 17502-510.

Rengasamy P (2010) Osmotic and ionic effects of various electrolytes on the growth of wheat. Soil Res 48:120-124

Resende Jr M, Muñoz P, Resende MDV, Garrick DJ, Fernando RL, Davis JM, Jokela EJ, Martin TA, Peter GF, Kirst M (2012) Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (Pinus taeda L.). Genetics 190:1503-1510

Reuzeau C, Frankard V, Hatzfeld Y, Sanz A, Van Camp W, Lejeune P, De Wilde C, Lievens K, de Wolf J, Vranken E (2006) Traitmill™ : a functional genomics platform for the phenotypic analysis of cereals. Plant Gen Res 4:20

Robbins MD, Massud Mohammed AT, Panthee DR, Gardner RG, Francis DM, Stevens MR (2010) Marker-assisted selection for coupling phase resistance to *Tomato spotted wilt virus* and *Phytophthora infestans* (Late Blight) in Tomato. Hort Sci 45:1424-1428

Saisho D, Takeda K (2011) Barley: emergence as a new research material of crop science. Plant Cell Physiol 52:724-727

Saito K, Hayano-Saito Y, Kuroki M, Sato Y (2010) Map-based cloning of the rice cold tolerance gene*Ctb1*. Plant Sci 179:97-102

Salvi S, Sponza G, Morgante M, Tomes D, Niu X, Fengler KA, Meeley R, Ananiev EV, Svitashev S, Bruggemann E (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc Natl Acad Sci 104:11376-11381

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552

Schaeffer ML, Harper LC, Gardiner JM, Andorf CM, Campbell DA, Cannon EKS, Sen TZ, Lawrence CJ (2011) MaizeGDB: curation and outreach go hand-in-hand. Database (Oxford) doi: 10.1093/database/bar022

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671-675

Schuster I (2011) Marker-assisted selection for quantitative traits. CBAB 11:50-55

Sheffield VC, Nishimura DY, Stone EM (1995) Novel approaches to linkage mapping. Curr Opin Genet Dev 5:335-341

Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A, Mungall CJ (2007) The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol 25:1251-1255

Smýkal P, Aubert G, Burstin J, Coyne CJ, Ellis NTH, Flavell AJ, Ford R, Hýbl M, Macas J, Neumann P, McPhee KE, Redden RJ, Rubiales D, Weller JL, Warkentin TD (2012) Pea (*Pisum sativum* L.) in the genomic era. Agronomy 2:74-115

Sorensen D, Gianola D (2002) Likelihood, Bayesian and MCMC methods in quantitative genetics. Springer doi: 10.1007/b98952

Steele K, Virk D, Kumar R, Prasad S, Witcombe J (2007) Field evaluation of upland rice lines selected for QTLs controlling root traits. Field Crops Res 101:180-186

Stein LD (2010) The case for cloud computing in genome informatics. Genome Biol 11:207

Swamy BPM, Vikram P, Dixit S, Ahmed H, Kumar A (2011) Meta-analysis of grain yield QTL identified during agricultural drought in grasses showed consensus. BMC Genomics 12:319

Tanabata T, Shibaya T, Hori K, Ebana K, Yano M (2012) SmartGrain: High-throughput phenotyping software for measuring seed shape through image analysis. Plant Physiol 160:1871-1880

Tavakkoli E, Rengasamy P, McDonald GK (2010) The response of barley to salinity stress differs between hydroponic and soil systems. Funct Plant Biol 37:621-633

Thomson MJ, Zhao K, Wright M, McNally KL, Rey J, Tung C, Reynolds A, Scheffler B, Eizenga G, McClung A, Kim H, Ismail AM, de Ocampo M, Mojica C, Reveche

MY, Dilla-Ermita CJ, Mauleon R, Leung H, Bustamante C, McCouch SR (2011) High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. Mol Breed 29:875-886 doi: 10.1007/s11032-011-9663-x

Trebbi D, Maccaferri M, de Heer P, Sorensen A, Giuliani S, Salvi S, Sanguineti MC, Massi A, van der Vossen EA, Tuberosa R (2011) High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). Theor Appl Genet doi: 10.1007/s00122-011-1607-7

Tuberosa R (2012) Phenotyping for drought tolerance of crops in the genomics era. Front Physio 3:347

Tucker SS, Craine JM, Nippert JB (2011) Physiological drought tolerance and the structuring of tallgrass prairie assemblages. Ecosphere 2:art48

Tung C, Zhao K, Wright MH, Ali ML, Jung J, Kimball J, Tyagi W, Thomson MJ, McNally K, Leung H, Kim H, Ahn S, Reynolds A, Scheffler B, Eizenga G, McClung A, Bustamante C, McCouch SR (2010) Development of a research platform for dissecting phenotype–genotype associations in rice (*Oryza* spp.). Rice 3:205-217 doi: 10.1007/s12284-010-9056-5

Valdar W, Flint J, Mott R (2006) Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. Genetics 172:1783-1797

Van Ommen B, Bouwman J, Dragsted LO, Drevon CA, Elliott R, de Groot P, Kaput J, Mathers JC, Müller M, Pepping F (2010) Challenges of molecular nutrition research 6: the nutritional phenotype database to store, share and evaluate nutritional systems biology studies. Genes & Nutrition 5:189-203

Varshney RK, Close TJ, Singh NK, Hoisington DA, Cook DR (2009) Orphan legume crops enter the genomics era! Curr Opin Plant Biol 12:202-210

Varshney RK, Glaszmann JC, Leung H, Ribaut JM (2010) More genomic resources for less-studied crops. Trends Biotechnol doi: 10.1016/j.tibtech.2010.06.007

Venuprasad R, Bool M, Quiatchon L, Sta Cruz M, Amante M, Atlin G (2012) A large-effect QTL for rice grain yield under upland drought stress on chromosome 1. Mol Breed 30:535-547

Venuprasad R, Impa S, Gowda R, Atlin G, Serraj R (2011) Rice near-isogenic-lines (NILs) contrasting for grain yield under lowland drought stress. Field Crops Res 123:38-46

Vikram P, Swamy B, Dixit S, Ahmed H, Cruz MT, Singh A, Kumar A (2011) qDTY1. 1, a major QTL for rice grain yield under reproductive-stage drought stress with a consistent effect in multiple elite genetic backgrounds. BMC Genetics 12:89

Vision T, Blake J, Lapp H, Mabee P, Westerfield M (2011) Similarity between semantic description sets: addressing needs beyond data integration. LISC2011 783

Wang L, Uilecan IV, Assadi AH, Kozmik CA, Spalding EP (2009) HYPOTrace: image analysis software for measuring hypocotyl growth and shape demonstrated on Arabidopsis seedlings undergoing photomorphogenesis. Plant Physiol 149:1632-1637

Wang Z, Yu C, Liu X, Liu S, Yin C, Liu L, Lei J, Jiang L, Yang C, Chen L (2012) Identification of Indica rice chromosome segments for the improvement of Japonica inbreds and hybrids. Theor Appl Genet 124:1351-1364

Weber JL, Broman KW (2001) 7 Genotyping for human whole-genome scans: Past, present, and future. Adv Genet 42:77-96

Weight C, Parnham D, Waites R (2007) TECHNICAL ADVANCE: LeafAnalyser: a computational method for rapid and large‐scale analyses of leaf shape variation. Plant J 53:578-586

White JW, Andrade-Sanchez P, Gore MA, Bronson KF, Coffelt TA, Conley MM, Feldmann KA, French AN, Heun JT, Hunsaker DJ (2012) Field-based phenomics for plant genetics research. Field Crops Res 133:101-112

Xu J, Zhao Q, Du P, Xu C, Wang B, Feng Q, Liu Q, Tang S, Gu M, Han B, Liang G (2010) Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (Oryza sativa L.) BMC Genomics 11:656. DOI: 10.1186/1471-2164-11-656

Yadav RS, Sehgal D, Vadez V (2011) Using genetic mapping and genomics approaches in understanding and improving drought tolerance in pearl millet. J Exp Bot 62:397-408

Yazdanbakhsh N, Fisahn J (2009) High throughput phenotyping of root growth dynamics, lateral root formation, root architecture and root hair development enabled by PlaRoM. Funct Plant Biol 36:938-946

Yin X, Struik P (2008) Applying modeling experiences from the past to shape crop systems biology: the need to converge crop physiology and functional genomics. New Phytol 179:629-642

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539-551

Yuan G, Luo Y, Sun X, Tang D (2004) Evaluation of a crop water stress index for detecting water stress in winter wheat in the North China Plain. Agric Water Manage 64:29-40

Zhang H, Zhao Q, Sun Z, Zhang C, Feng Q, Tang S, Liang G, Gu M, Han B, Liu Q (2011) Development and high-throughput genotyping of substitution lines carrying the chromosome segments of indica 9311 in the background of japonica Nipponbare. J Genet Genomics 38:603-611

Zhang J, Li C, Wu C, Xiong L, Chen G, Zhang Q, Wang S (2006) RMD: a rice mutant database for functional analysis of the rice genome. Nucleic Acids Res 34:D745-D748

Zhao K, Tung C, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, Bustamante CD, McCouch SR (2011a) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza* sativa. Nat Com 2:467

Zhao Y, Gowda M, Liu W, Wurschum T, Maurer HP, Longin FH, Ranc N, Reif JC (2011b) Accuracy of genomic selection in European maize elite breeding populations. Theor Appl Genet. doi: 10.1007/s00122-011-1745-y

Zhong S, Dekkers JCM, Fernando RL, Jannink JL (2009) Factors affecting accuracy from genotypic selection in populations derived from multiple inbred lines: A barley case study. Genetics 182:355-364

Zhou H, Liu Q, Li J, Jiang D, Zhou L, Wu P, Lu S, Li F, Zhu L, Liu Z (2012) Photoperiod-and thermo-sensitive genic male sterility in rice are caused by a point mutation in a novel noncoding RNA that produces a small RNA. Cell Res 22:649-660

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. The Plant Genome 1:5-20

Zhu D, Deng XW (2012) A non-coding RNA locus mediates environment-conditioned male sterility in rice. Cell Res 22:791-792

Zhu J, Ingram Pa, Benfey PN, Elich T (2011b) From lab to field, new approaches to phenotyping root system architecture. Curr Opin Plant Biol 14:310-317

Zhu J, Wang X, Sun C, Zhu X, Li M, Zhang G, Tian Y, Wang Z (2011a) Mapping of QTL Associated with Drought Tolerance in a Semi-Automobile Rain Shelter in Maize (Zea mays L.). Agric Sci China 10:987-996

Zia S, Romano G, Spreer W, Sanchez C, Cairns J, Araus J, Müller J (2012) Infrared Thermal Imaging as a Rapid Tool for Identifying Water ‐ Stress Tolerant Maize Genotypes of Different Phenology. J Agron Crop Sci doi:10.1111/j.1439-037X.2012.00537.x

CHAPTER 3:

IONOMIC PROFILING IN A PANEL OF DIVERSE *Oryza sativa* ACCESIONS IS ASSOCIATED WITH NATURAL GENETIC VARIATION AS REVEALED BY GENOME-WIDE ASSOCIATION MAPPING

ABSTRACT

Nutrient acquisition and exclusion of toxic elements is of paramount importance to all plant life, and of particular interest to crop species that form the basis of the human food supply. Ionomic profiling offers a way of simultaneously analyzing the entire mineral element composition of tissues within an organism. Rice is an ideal system for exploring the functional genomics of the ionome because of its diverse genetic resources, wide array of genomics tools, ability to grow in diverse soils and environments, and importance as a staple crop. In this study, we screened 394 diverse *O. sativa* accessions under controlled conditions in a specially designed nutrient solution to survey uptake and partitioning of 24 elements in roots and shoots of 6-week old seedlings. We conducted Genome-Wide Association (GWA) analysis using SNPs from the recently developed High Density Rice Array (HDRA) to associate natural genetic variation with root and shoot ionomic content across and within the major subpopulations of domesticated Asian rice. Distinct networks of ions were associated with ionomic content of roots versus shoots, and among the different sub-populations in our diversity panel. A major peak in the analysis highlights the potential role of *OsMOT1* in governing shoot molybdenum content. Likewise we

report on the co-localization of a small heat shock protein (*OsSLT1*) with a major peak for sodium accumulation in the *aus* subpopulation. *SLT1* was previously implicated in sodium homeostasis in *Arabidopsis* and Tobacco via yeast mutant complementation, and here we report that haplotype variation at this gene in rice is significantly associated with shoot Na content and shoot Na/K ratio, both important measures of yield under salt stress. We also identified two individual accessions (NSF_TV_11 and NSF_TV_171) that preferentially discouraged translocation of Cd to the shoots. The *aus* subpopulation in particular is highlighted as a potentially valuable donor of genetic variation for the repartitioning of Cu, Fe, Mg, Mo, P, Si, Sr, and Zn from the root to the shoot.

INTRODUCTION

The processing and acquisition of matter is of seminal importance to life on earth, and is of particular concern for plants whose sedentary life cycle limits the organism to acquiring the necessary materials from a confined radius of soil surrounding the plant. The acquired mineral elements are critical to the ability of nucleic acids, amino acids and carbohydrates to perform their biological functions; 14 mineral elements have been identified as essential to all plant life. Six of these are commonly known as macronutrients and are required in relatively large amounts (>0.1% of dry mass). These consist of nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), and sulfur (S) (Maathuis 2009). A second group of elements are no less essential, but required in much smaller, trace amounts. These micronutrients include iron (Fe), boron (B), choride (Cl), copper (Cu), manganese (Mn),

molybdenum (Mo), zinc (Zn), and nickel (Ni) (Hänsch and Mendel 2009). In addition to the essential macro- and micro-nutrients that plants need to survive, plants have evolved to exclude or avoid the acquisition of toxic elements that are counter-productive to biological processes. Yet many toxic elements share analogous chemical properties to essential nutrients, making them difficult to discriminate. Such elements include arsenic (As), cadmium (Cd), chromium (Cr), lithium (Li), sodium (Na), lead (Pb), strontium (Sr) and aluminium (Al). While in many cases natural deposits of some of these elements are a source of toxicity to plants, others come from anthropogenic contamination of soils. Strontium, as one example, is one of the most hazardous radioactive byproducts of nuclear fission and plants are much more adept at absorbing it from the environment than are animals (Soetan et al. 2010). While typically neither toxic nor essential, there are other elements such as cobalt (Co), iodine (I), rubidium (Rb), selenium (Se) and silicon (Si) which play important roles in the plant's adaptive responses to their environment or are otherwise necessary for human nutrition and thus are important selection targets for plant breeders (Pilon-Smits et al. 2009, Soetan et al. 2010). The appropriate acquisition of essential and beneficial elements, coupled with the necessity to avoid toxic or harmful elements, is further complicated for plants due to the highly variable nature of the distribution of these elements in the environment over space and time. Their distribution in soil is a result of stochastic geological, evolutionary and environmental factors including weather, climate, erosion, physiochemical properties, and pH. As such, in order to sustain growth, plants have had to develop a suite of adaptive and environmentally

responsive mechanisms for the acquisition and internal distribution of mineral elements (Williams and Salt 2009).

The emergence of 'ionomics' represents a high throughput approach to the phenotyping of plant populations and offers an important foundation that will allow us to begin to understand the complex physiological networks that plants have evolved to cope with the challenges of acquiring appropriate nutrition. Simply put, the plant ionome is "the mineral nutrient and trace element composition of an organism and represents the inorganic component of cellular and organismal systems" (Lahner et al. 2003, Salt et al. 2008). Understanding the genetic architecture and underlying functional genomics of the ionome in major crop species will provide important new insights about the range of phenotypic variation within and between crop species and the complex trade-offs that define opportunities to optimize productivity in the face of soil and nutrient constraints, climate change, and the need to improve the long-term sustainability of crop production systems.

For many decades, inroads have been made in understanding the underlying molecular mechanisms of nutrient uptake and distribution on an element-by-element basis (see Williams and Salt (2009) for a survey of several reviews on this topic published in a special issue of Current Opinion in Plant Biology). Over the last decade, new tools and technologies have enhanced the power and resolution of genome-wide approaches for both forward and reverse genetic analysis, enabling the scientific community to begin to dissect the genetic architecture of the ionome as a whole. To date, the most

extensive effort to understand the functional genomics governing the ionome has been done using mutagenized populations of *Arabidposis thaliana* (Lahner et al. 2003), *Saccharomyces cerevisiae* (Eide et al. 2005), and the model legume *Lotus japonicas* (Chen et al. 2009). Since that time there have been also several attempts to characterize natural variation in plant ionomic profiles, notably in *Lotus japonicas* (Chen et al. 2009, SANCHEZ et al. 2011), *Arabidopsis thaliana* (Atwell et al. 2010, Buescher et al. 2010, Ghandilyan et al. 2009, Rus et al. 2006, Vreugdenhil et al. 2004), *Mimulus guttatus* (Seep Monkeyflower) (Lowry et al. 2012), *Oryza sativa* (Li et al. 2013, Norton et al. 2010, Norton et al. 2012, Norton et al. 2009), and *Zea mays* (Baxter et al. 2013), and broadly across 138 distinct families of terrestrial plants (Watanabe et al. 2007).

Current efforts to associate ionomic profiles with genetic variation at the population level in *Arabidopsis* have identified at least six genes that govern mineral nutrient homeostasis for metal ions (Rampey et al. 2006), sodium (Rus et al. 2006), molybdenum (Baxter et al. 2008), iron and cobalt (Morrissey et al. 2009), cadmium (Chao et al. 2012), and the ionome more generally via a root suberin protein (Baxter et al. 2009). Four of these genes were implicated as natural variants among diverse *Arabidopsis* accessions, while the other two (Rampey et al. 2006 and Baxter et al. 2009) were only identified as induced mutations in the Columbia (Col-0) wild type Arabidopsis genetic background.

Published reports of whole ionome profiling and the association of genotype with ionomics phenotypes in domesticated Asian rice (*Oryza sativa*) thus far have been limited to QTL mapping in a single population of recombinant inbred lines (RILs) derived from a cross between a traditional upland *tropical japonica* variety (Azucena) and a lowland *aus*-like *indica* variety (Bala), and have focused specifically on the leaf and grain ionome, in most cases contextualized by a search for the genes /QTLs underlying arsenic toxicity (Norton et al. 2010, Norton et al. 2012, Norton et al. 2009). Recently two studies surveyed the grain ionomic profiles of contemporary paddy rice as well as rice landraces sampled from a limited geographic region in China (Li et al. 2013, Zeng et al. 2010). The attempt of Zeng et al. (2010) to associate allelic variation of 20 SSRs in Chinese landraces to phenotypic variation of the unpolished grain ionome demonstrated a clear link between genetic variation and ionomic profile phenotypes and emphasized the need to use more sophisticated genetic resources and marker technologies to explore this complex relationship between genotype and phenotype.

O. sativa is arguably one of the most important food crops in the world, providing up to 76% of the caloric intake for many populations in South East Asia and roughly 1/5 of human caloric intake worldwide (Fitzgerald et al. 2009). In addition to being a major target species for both private and public sector breeding efforts due to it's role as a staple food in the human diet, rice comparative genomics also allows researchers to build on the wealth of rice sequence information (Goff et al. 2002, Yu et al. 2002), mapping and germplasm resources (Agrama et al. 2010, Huang et al. 2010, McCouch

et al. 2012, McCouch et al. 2010, McNally et al. 2009, Tung et al. 2010, Zhao et al. 2011) and publicly available databases for the benefit of related cereal grains (Sakai et al. 2013, Youens-Clark et al. 2011). As a model system, these resources make rice a powerful starting point for investigating the genetic basis for complex trait variation.

The widely divergent phenotypic variation present in rice germplasm has led to a historical classification of rice into two varietal groups, *Indica* and *Japonica* (indicated in this manuscript in italics and beginning with a capital letter), which has since been confirmed using molecular tools (Huang et al. 2012, Kovach and McCouch 2008, Molina et al. 2011, Oka 1988). The population sub-structure of these clades is divided into 5 distinct subgroups, referred to as subpopulations (indicated in this manuscript in italics with a lower case letter). These include *indica, aus, temperate japonica, tropical japonica* and *aromatic* [*Group V*] (Garris et al. 2005, Zhao et al. 2011).

While documentation of natural variation for elemental uptake, sequestration, utilization, and mobilization in rice has begun, little is known about the identity and regulation of the genes involved. Combining high-quality rice genomic resources (Tung et al. 2010) with sophisticated and carefully designed phenotyping approaches and statistical analysis should yield (Cobb et al. 2013) a better and more detailed understanding of the genetic architecture of mineral element homeostasis. This in turn will empower the effective integration of novel genetics for the breeding of nutrient fortification, input use efficiency, and abiotic stress tolerance into modern, high-yielding, resource-use efficient cultivars.

Furthermore, from a purely genetics standpoint the ionome is an interesting phenotype to work with, as it represents a correlated network of related phenotypes rather than one physiological/morphological metric. Understanding the genetic architecture of the overall phenotypic network is helpful when trying to manipulate complex but poorly interrogated traits such as nutrient acquisition when only one or a few mineral nutrients are considered at a time. In many cases, interesting and nonintuitive relationships emerge when characterizing the distribution of the broader ionome (Eide et al. 2005, Lahner et al. 2003). This principle is ironically juxtaposed with the use of single marker association analysis as a tool to identify genomic regions with phenotypic impacts on the rice ionome. The use of multi-variate analysis techniques will be needed to characterize the relationships among traits and to begin to explore the trade offs that occur in the ionome as a whole when the uptake or distribution of one element is perturbed due to either genetic variation in the plant or changes in availability of ions in the environment.

It's worthy to note an important limitation of the ICP-OES analytical technology as applied to the ionome. Nitrogen and chlorine are both essential nutrients to plants that can't be surveyed by the optical emission technology due to the necessity to digest samples in nitric and perchloric acid. Chlorine is essential and needed in only very trace amounts, but Nitrogen is the single most important elemental determinant of agricultural productivity. These two elements were not included in the present study

due to throughput limitations, and more intensive research into their homeostatic maintenance is warranted; particularly as it intersects with nitrogen use efficiency.

The objectives of this study were: [1] to define a controlled hydroponic-based phenotyping platform that allowed precise and accurate determination of root and shoot accumulation of 24 elements of agronomic or biological importance to rice, [2] to measure root and shoot nutrient density using inductively coupled plasma-optical emission spectroscopy [ICP-OES] in a publicly available rice association mapping population. This population, which is referred to as the Rice Diversity Panel 1 (RDP1), consists of a collection of 410 diverse domestic rice varieties collected from 79 countries (Eizenga et al. 2013, Ali et al. 2011, Tung et al. 2010) that has been genotyped using a 44,000 Affymetrix custom SNP array (Zhao et al. 2011) and more recently, a high-density Affymetrix custom SNP array (HDRA; High-Density Rice Array), providing a resolution of one SNP every 3-5 kilobases (McCouch S., Cornell University, personal communication), [3]to associate the phenotypic variation for mineral density with genetic polymorphisms and identify genomic regions associated with uptake, sequestration and partitioning of each element using Bayesian estimates of phenotypic means and single SNP linear mixed model analysis, and [4] to identify candidate genes underlying GWA peaks and characterize their haplotype signatures using whole-genome re-sequencing data from 125 rice lines (riceSNP.org; Wright, M. and McCouch S., Cornell University, personal communication) in order to gain insight into the genetic architecture of the phenotypic variation they control.

Two of the primary discoveries immediately suggested by GWA analysis in this study concern the effects of natural genetic variation at key candidate gene loci for shoot sodium and shoot molybdenum accumulation. Molybdenum is essential for plants but needed in only trace amounts. Only four proteins are known to require molybdenum (nitrate reductase, sulfite oxidase, aldehyde oxidase, and xanthine dehydrogenase; (Hänsch and Mendel 2009) but these proteins are involved in important physiological processes such as the assimilation and metabolism of both nitrogen and sulfur, phytohormone biosynthesis, and stress responses. The role of molybdenum in nitrate assimilation is of particular interest as the processing and acquisition of nitrogen is paramount to plant growth and reproduction, while at the same time nitrogen use efficiency is a major breeding objective. Molybdenum is critical for the proper functioning of nitrate reductase, which catalyzes the first and rate limiting step of nitrogen assimilation (Schwarz and Mendel 2006). Furthermore, the proper functioning of aldehyde oxidase is critical for abscisic acid biosynthesis, and sulfite oxidase is instrumental in protecting plants from the toxic effects of sulfite, the primary component of acid rain (Hänsch and Mendel 2009). Complexation of molybdenum with molybdopterin (a metal-containing pterin compound) is necessary to facilitate biological activity, and collectively the molybdenum-molybdopterin complex is referred to as molybdenum cofactor (moco). Molybdenum metabolism interacts both upstream and downstream of moco biosynthesis with pathways governing iron, copper, and sulfur metabolism, as well, and together represent crosstalk between ionomic networks within the plant (Hänsch and Mendel 2009, Kuper et al. 2004). Our study is the first investigation to describe a molybdate

transporter in rice, to document the effect of natural variation at that locus on shoot molybdate concentration, and to propose a mechanism for how allelic variation at that locus contributes to whole shoot molybdate content.

Sodium on the other hand has been widely studied due to the toxic effects it exhibits on plant growth and the abundance of saline/sodic soils in which rice is grown. Low levels of sodium are beneficial to plants as the sodium cation can act as an important osmolyte and can replace potassium as a cofactor in certain chemical reactions (Pilon-Smits et al. 2009). Sodium has also been shown to be an important component of CAM/C4 photosynthetic pathways (Ohnishi et al. 1990). Sodium transport and metabolism is probably best known, however, for the exclusion, or vacuolar sequestration of, sodium in the event of salt stress (Munns and Tester 2008, Zhu 2000). Relative to the well-documented role of the high-affinity K+ transporters (HKT's) in the uptake and repartitioning of sodium, the role of small heat shock proteins (sHSPs) such as *SLT1* has received little attention (Antoine et al. 2005, Garciadeblas et al. 2003, Matsumoto et al. 2001). To the author's knowledge, this is the first report of natural variation for shoot sodium accumulation implicating an *SLT1* ortholog in plant sodium homeostasis , and this is the first report suggesting the involvement of *OsSLT1* in sodium metabolism and shoot sodium accumulation in rice. Additionally we report evidence of the possible importance of *OsSLT1* not only in sodium homeostasis, but also in sodium exclusion from the shoot under salt stressed conditions.

Given the breadth of our elemental analysis (24 chemical species), the extent of the genetic variation in the rice diversity panel, the density of the SNP coverage available on this panel, the depth of re-sequencing data available on the 125 lines used for sequence analysis and haplotype reconstruction, the precision of our phenotyping platform, and the sophistication of our Bayesian estimation of phenotypic parameters, to our knowledge this is the most extensive survey of genotype-phentoype relationships in a single species for ionomic variation published to date.

RESULTS AND DISCUSSION

Precision and accuracy of the phenotyping platform

Three hundred ninety-four diverse accessions from the NSFTV rice diversity panel (Eizenga et al. 2013, Liakat Ali et al. 2011, Tung et al. 2010, Zhao et al. 2011) were grown hydroponically to 47 days post-germination in a standard nutrient solution modified to provide all of the macro- and micronutrients necessary for plant growth as well as sub-toxic levels of heavy metals and other ions that would be non-standard in a nutrient solution but are of biological or agronomic interest. Macronutrient levels were also adjusted as to minimize competition for uptake between different macronutrients but were still at adequate levels so as not to induce nutrient deficiencies. pH of the nutrient solution was maintained at 6.0 for the duration of the experiments. This is different from the pH of a rice paddy (typically 7.0) in order to 1) ameliorate the chemical sequestration of nutrients as the pH becomes more acidic over the course of 24 hours due to ammonium uptake; 2) to ensure the general availability of all ions in the analysis, some of which lose availability at higher pH values, and 3) reduce the

amount of potassium hydroxide needed to maintain the pH of the nutrient solutions over time. Even at pH 6.0 the additional potassium added over the course of the experiment was enough to saturate most genotypes and reduce the variance of potassium shoot concentration compared to the other ionomic phenotypes (Supplemental Figure 3.22).

By 47 days post-germination, all of the accessions in the population were in the vegetative growth stage and had an average root dry-weight biomass of 172 milligrams. This was critical as the ability to precisely detect low levels of trace elements with an inductively coupled plasma optical emission spectrometer (ICP-OES) depended on a minimum atomic signal in the sample tube. Using between 100 and 200 mg of dry weight root tissue provided enough signal from the ionized sample to reproducibly ascertain the lowest concentrations of elements in the rice samples (Table 3.1). To achieve the desired levels of precision for ICP-OES analysis of root tissue, the entire shoot biomass growth (average of 1.56g) often exceeded the amount of tissue that could be analyzed in a single sample. As a result, subsamples of 200 mg of homogenized shoot tissue were used to estimate whole shoot elemental concentrations.

In addition to small sample volumes, one other significant source of variation in an ICP-based elemental analysis pipeline is the dilution volume used to calculate the concentration of each element within a sample. Small variations in the delivery volume of the acid solution used to dilute the samples as well as evaporation after

dilution has taken place could change the final dilution volume. In an effort to improve dilution accuracy, a known volume and concentration of yttrium was added to each sample. Yttrium is not absorbed by plant tissue and thus is used to determine the total volume of each sample. The yttrium concentration data produced by the ICP is compared to the amount of yttrium that was added to calculate an individualized dilution factor for every sample in the analysis. This had the added benefit of accounting for drift in the ICP measurements that inevitably occurs over runs with large sample numbers and is typically corrected only when the occasional quality control solution is inserted between experimental samples.

Element	Avg %RSD	Lower Limit of Detection (in ppm)
Arsenic (As)	**3.94**	**0.0383**
Boron (B)	**5.10**	**0.1463**
Calcium (Ca)	**2.84**	**0.2913**
Cadmium (Cd)	**1.54**	**0.0044**
Cobalt (Co)	**4.87**	**0.0028**
Chromium (Cr)	**2.70**	**0.0219**
Copper (Cu)	**2.52**	**0.0053**
Iron (Fe)	**3.36**	**0.2399**
Iodine (I)	**74.14**	**0.3569**
Potassium (K)	**3.97**	**6.864**
Lithium (Li)	**88.80**	**0.0287**
Magnesium (Mg)	**3.00**	**1.219**
Manganese (Mn)	**3.63**	**0.0009**
Molybdenum (Mo)	**3.36**	**0.0034**
Sodium (Na)	**3.75**	**0.788**
Nickel (Ni)	**55.88**	**0.0028**
Phosphorus (P)	**1.81**	**2.794**
Lead (Pb)	**48.98**	**0.0077**
Rubidium (Rb)	**121.12**	**0.2264**
Sulfur (S)	**2.48**	**0.5086**
Selenium (Se)	**2.29**	**0.0142**
Silicon (Si)	**2.62**	**0.0886**
Strontium (Sr)	**2.45**	**0.0008**
Zinc (Zn)	**1.96**	**0.0651**

Table 3.1: Percent relative standard deviations and lower limits of detection for the 24 elements in the analysis. Table depicts the respective precision (%Relative Standard Deviation--%RSD) and accuracy (Lower Detection Limit) of our phenotyping platform. %RSD values reflect an average across all samples. %RSD for each sample was calculated as 100*(Standard Deviation/Mean) of three repeated measurements. Lower limits of detection (LOD; expressed as ppm) for each element were calculated as 3*(Standard Deviation of a blank sample).

Once the means and %RSD's were known for each sample, a Bayesian modeling approach was used to account for the experimental covariates (such as tub, plant position within the tub, digestion batch etc…) and estimate a mean phenotypic value (50^{th} percentile of the posterior distribution) for each line in the population. To extract robust estimates of means from replicated measurements of ion concentration, we constructed a hierarchical Bayesian model (Gelman and Hill 2007). This model considers all phenotypes at once, estimates covariance among them, and takes into account the uncertainty in all parameters to construct marginal distributions of all variables of interest. For the purposes of this study, we are primarily interested in the portion of the phenotypic attribute of each rice line that is due to its genetic make-up and our modeling follows very closely the ideas presented in (Greenberg et al. 2011). We assume that covariances come from inverse-Wishart distributions with vague diagonal priors. We have three levels of replication within which all the phenotypes are simultaneously modeled: 1) individual observations are nested within lines, which are nested 2) within subpopulations, with 3) an overall mean for the entire diversity panel as the final level. Owing largely to the significant amount of tissue used for the analyses, the controlled conditions under which the plants were grown, and the sophistication of modeling experimental covariates to determine line means, the heritability for most of the ionomic phenotypes in this study is relatively high (≥ 0.4) (Figure 3.1). This results from a reduction in type I error due to the correction for covariates such as tub identity, plant position, initials of the harvester, and the ICP digestion batch. While critical to reducing noise in the data set it does little to correct for type II errors, which can only come with increased population size (see Chapter 4).

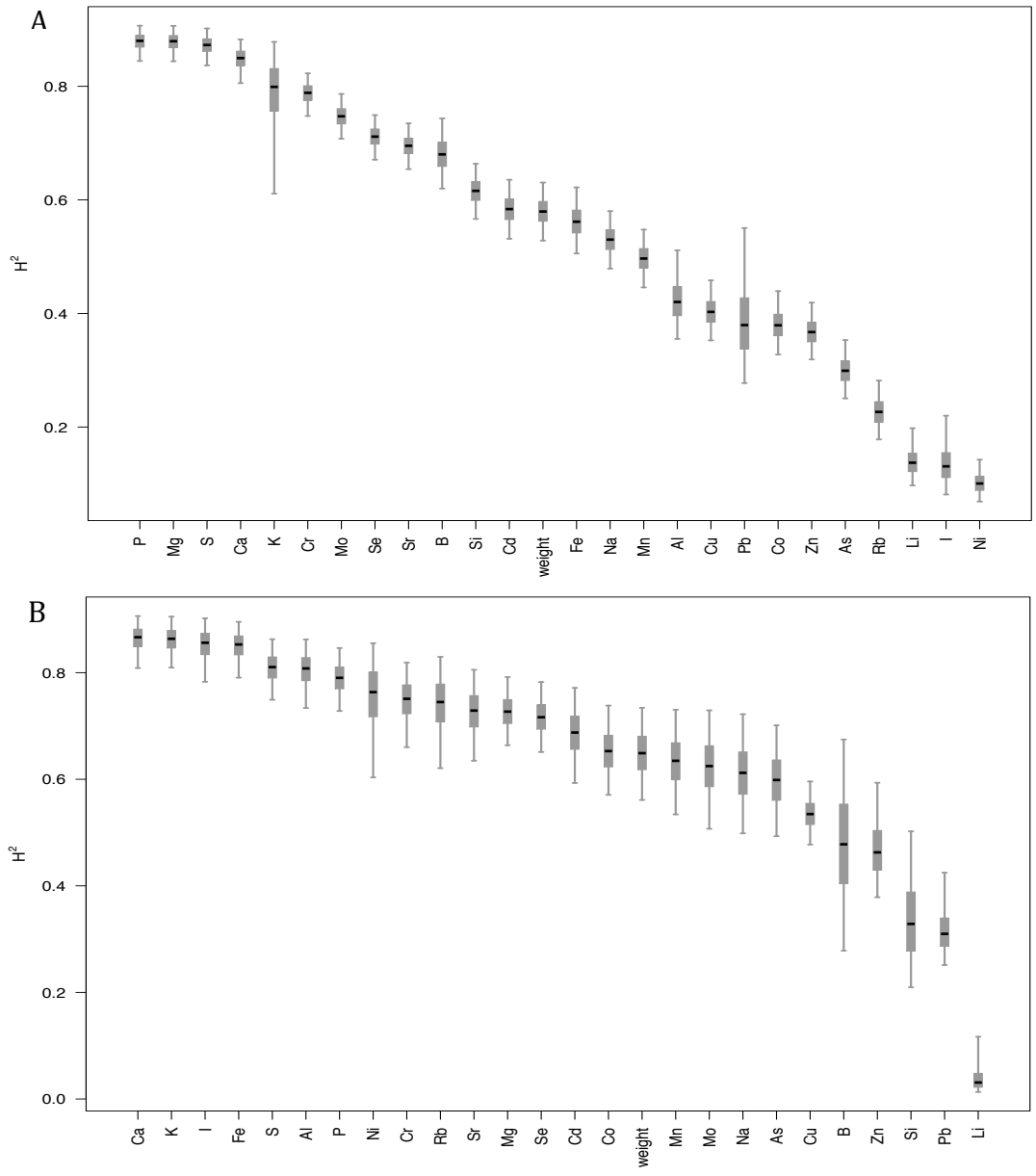


Figure 3.1: Respective heritability for each ionomic phenotype. Bayesian estimates of broad-sense heritability (H^2 = Var(among lines)/[Var(among lines) + Var(among replicates)]) of A) whole shoot phenotypes and B) whole root phenotypes. Black bars indicate the 50th percentile of the posterior distribution, and whiskers indicate the 95% credible interval of the posterior distribution.

Relationships among phenotypes within and across subpopulation structure

It is of interest to note that despite extensive control over the environmental variation and subsequent modeling of remaining extemporaneous sources of phenotypic variation, there is still a large range of heritability manifested across the rice ionome. Those elements with heritability below 0.3 (namely Li, I, Rb, Ni, and Pb) correspond to the elements whose percent relative standard deviations (%RSD; 88.80, 74.14, 121.12, 55.88, and 48.98 respectively) were high enough to preclude a precise estimation of their concentration. As such it is reasonable that this lack of precision would contribute to a larger variance within replicates of the same lines. For the remaining elements, the precision of phenotypic estimation was quite high, and as such the range of heritability among them could represent the relative complexity of the genetic architecture governing their uptake into the root and subsequent partitioning to the shoot tissue. For most elements of interest in rice, accumulation in the edible portions of the plant is of greater interest than that observed in roots or shoots. Since partitioning from the root to the shoot, however, is the first step in translocation to the grain, those elements with high shoot heritability have crossed the first hurdle toward breeding efforts to either increase (e.g. Fe bio-fortification) or decrease (e.g. Cd exclusion) their levels in the developing grain endosperm. It also stands to reason that heritability in the roots would be higher than in the shoots because partitioning decisions throughout the life cycle of the plant introduce more variation. Elemental species are often stored in different biological constructs that have different physiological properties that govern how they interact with the

biochemistry of an organism. Additionally whole shoot tissue in rice comprises both young and old tillers alike which are likely to exhibit differing ionomic profiles.

The relationship between root and shoot mineral concentration was highly variable across a wide range of genetic variation as well as across the various chemical species that comprise the ionome. For some elements, there was a positive correlation between root and shoot accumulation (e.g., Cr, Mo, Cd), others show a negative correlation (e.g., Fe, B, Ca), and many show no significant correlation between roots and shoots (Figure 3.2).

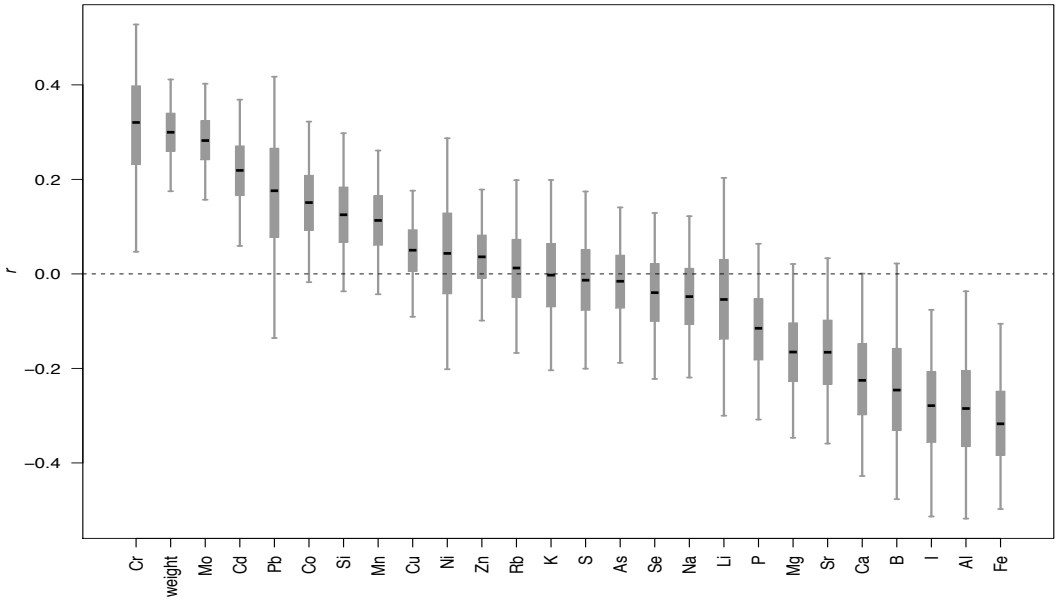


Figure 3.2 Bayesian estimation of correlations between root and shoot values for each phenotype. Values near r=0 indicate there is no relationship between root and shoot values. Positive values imply a positive correlation while negative values imply a physiological preference for distribution in one organ or the other. Error bars represent the 95% credible interval of the posterior distribution and black bars indicate the value of the 50th percentile of the same distribution.

Root/shoot ratios are another important metric for determining how different mineral elements are partitioned within a plant. Figure 3.3 illustrates the distribution of mean root/shoot ratios across genetic supopulations for each element in the rice ionome. The *indica, aus, tropical japonica* and *temperate japonica* subpopulations were all represented by >50 individuals, while ionomic data was available for only 12 samples in the *aromatic* (*Group V*) subpopulation. The *aromatic* subpopulation was therefore not considered further, due to lack of statistical power to make inferences.

All of the heavy metals (Cd, Fe, Ni, Pb, Sr, and Zn) were more abundant in root tissue, while elements known to be associated with shoot/leaf physiology (K, Mg, and Si) were more abundant in the shoot tissue. For many cations and in particular divalent cations (namely Cd, Co, Cu, Fe, Li, Na, Ni, Pb, Sr, and Zn), some of the high concentrations reported in root tissue can be attributed in part to binding and adsorption in the root cell wall because of the negative charge within plant cell walls creates a cation exchange matrix. This remains true even after careful rinsing and desorption of the root systems prior to ICP analysis.

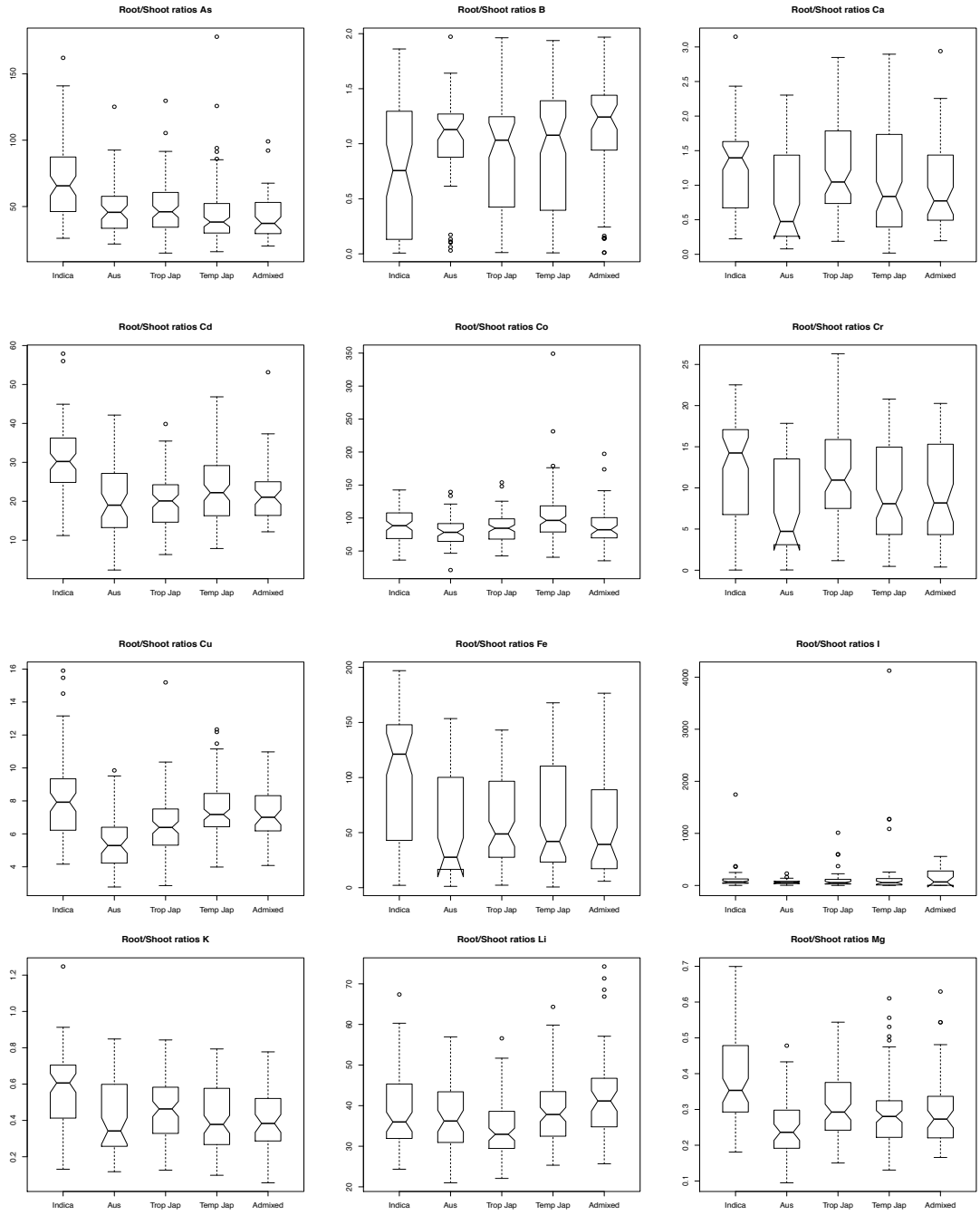


Figure 3.3 Box and whisker plots highlighting the differences of root/shoot ratio of each ionomic phenotype across the genetic substructure of rice.

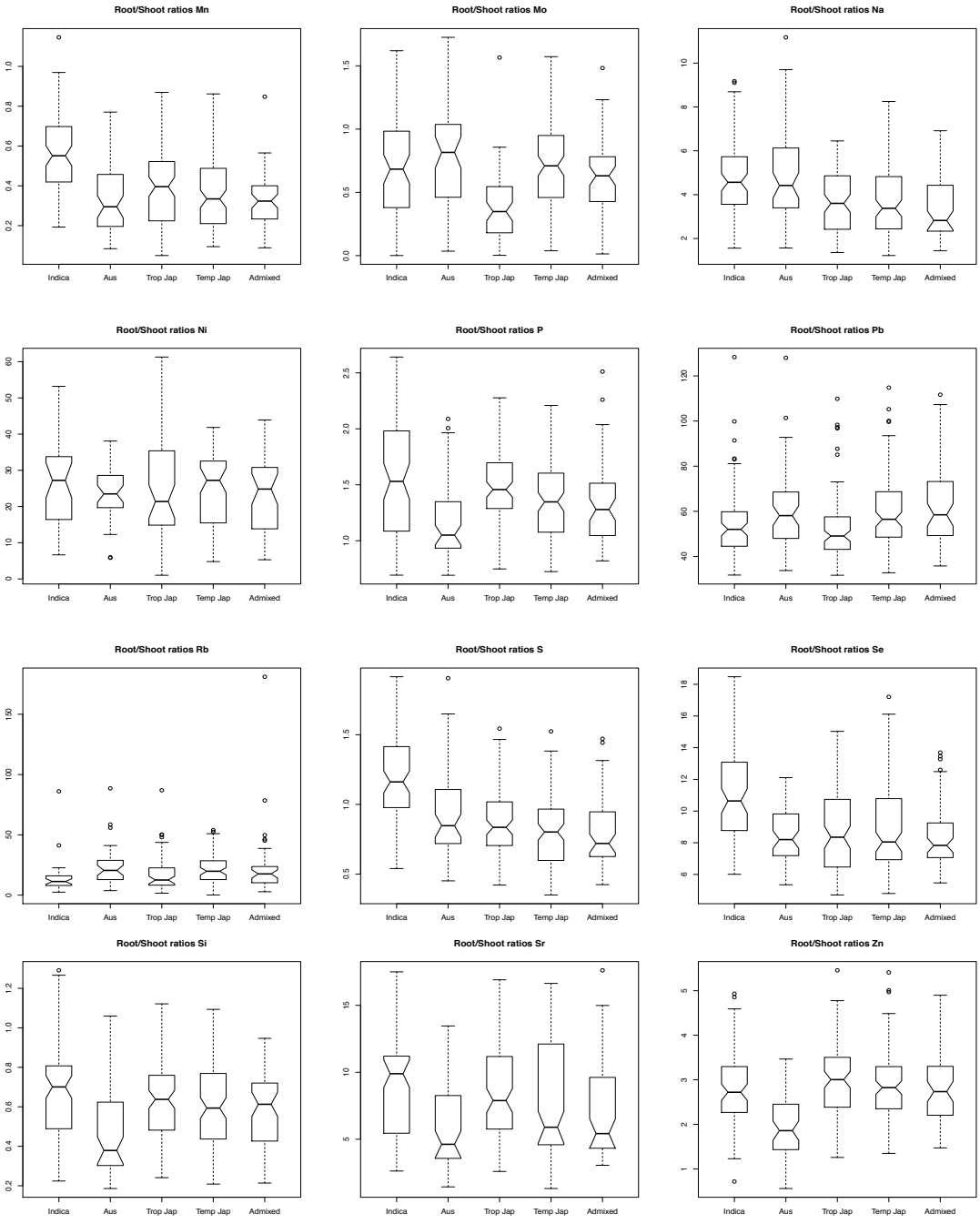


Figure 3.3 (continued)

Of interest to breeders and geneticists is the observation that there are a few outliers in one subpopulation or another for virtually every element (Fig. 3.3). These individuals represent transgressive segregants within a subpopulation and may contain key introgressions from another subpopulation that govern elemental re-distribution within the plant. For example, two *indica* varieties (NSF_TV_11 from Bangladesh and NSF_TV_171 from China) have root/shoot ratios for Cd that are almost two orders of magnitude greater than the same parameter values for *indica* varieties as a whole, indicating a clear preference for shoot cadmium exclusion. Further, NSF_TV_11 exhibited transgressive behavior relative to it's partitioning of manganese (a known chemical analog to cadmium) to root tissue. In fact, it was the only variety with a manganese root/shoot ratio greater than one, indicating that it is the only line that preferentially sequesters Mn in the roots.

Differences between subpopulation means were significant for every element except B and Ni and Tukey's multiple comparison test showed only Na exhibited a characteristic *Indica/Japonica* split (Table 3.2). Much more common was for the *indica* subpopulation to be significantly different than all the others. This may be due to the fact that the indica gene pool is far more diverse than the others and therefore is more prone to have transgressive genotypes. For many elements (Ca, K, Mn, Mo, P, S, and Si) the root/shoot ratio value spans 1.0 with some genotypes with a value higher than one and others with a value lower than one. A value of one is indicative of ions being distributed evenly on a per gram basis between roots and shoots. Since values above 1.0 indicate preferential sequestering of ions to the roots, and values below 1.0

indicate preferential reapportioning of ions to the shoots, these seven elements exhibit unique distributions with means around 1.0.

Of particular note are Ca and S because the differences in means for root/shoot ratios for these elements in the different subpopulations are both significant and span the value 1.0. In both cases the subpopulation means for *aus* and *indica* are significantly different where the *aus* subpopulation has a value of less than one showing disproportionate repartitioning to shoot tissue and the *indica* subpopulation mean is greater than one highlighting a preference of *indica* to keep higher concentrations of Ca and S in the root. This hypothesizes that some subpopulations may partition ionic species differently and underscores the genetic control of ion distribution with in the plant. It is unclear what this may mean for calcium and/or sulfur metabolism, but highlights the potential role of subpopulation-specific introgressions on the re-partitioning of sulfur within a variety.

Another interesting finding is that the *aus* subpopulation in particular exhibits a significant preference for partitioning of Ca, Cu, Si, Sr, and Zn to the shoot, relative to other subpopulations, and the *indica* subpopulation shows a significant preference for partitioning Cd, Fe, K, Mg, Mn, S, and Se to the root tissue. This implicates germplasm in the *Indica* clade as a potentially important source of genes controlling the redistribution of these elements to the above ground plant tissues.

Phenotype	*indica*	*aus*	*trop jap*	*temp jap*	ANOVA p-val	% Var exp by Subpop
As	a	b	b	b	< 0.00001	14%
B	a	a	a	a	0.05426	NS
Ca	a	b	a	ab	0.000483	5%
Cd	a	b	b	b	< 0.00001	20%
Co	a	a	a	b	< 0.00001	7%
Cr	a	bc	ad	cd	< 0.00001	7%
Cu	a	b	c	a	< 0.00001	17%
Fe	a	b	b	b	< 0.00001	12%
I	a	a	a	a	0.3557	0.10%
K	a	b	b	b	< 0.00001	10%
Li	a	ab	b	a	0.0014	4%
Mg	a	b	c	bc	< 0.00001	15%
Mn	a	b	b	b	< 0.00001	19%
Mo	a	a	b	a	< 0.00001	12%
Na	a	a	b	b	< 0.00001	10%
Ni	a	a	a	a	0.906	NS
P	a	b	ac	c	< 0.00001	10%
Pb	ab	a	b	a	0.003218	3%
Rb	a	b	b	b	0.02671	3%
S	a	b	b	b	< 0.00001	20%
Se	a	b	b	b	< 0.00001	10%
Si	a	b	a	a	< 0.00001	7%
Sr	a	b	a	a	< 0.00001	7%
Zn	a	b	a	a	< 0.00001	18%

Table 3.2 Results of Tukey's Honest Signficant Difference test for multiple comparisons of root/shoot ratios. Significantly different groups are indicated with different letters. P-values given and percent variance explained indicate results from two-way analysis of variance.

Relationships across space and time within a plant are one way to understand a phenotypic network, but often relationships between the phenotypes themselves within a plant tissue can be equally informative. Scaled principle component analyses are shown as biplots (Gabriel 1971) in Figures 3.4 and 3.5 (and Supplemental Figures 3.1 and 3.2). These analyses were conducted in an effort to begin to describe multivariate relationships among the elements of the ionome within roots and shoots. The length of each of the eigenvectors is proportional to the phenotypic variance for that element, and the angle between eigenvectors indicates a relationship between two phenotypes. Parallel unidirectional vectors indicate a strong positive correlation; parallel bidirectional vectors indicate a strong negative correlation; and perpendicular vectors are consistent with a lack of correlation between two variables. Similar descriptions of elemental interconnectedness for mutant populations have been described for *Arabidopsis* (Lahner et al. 2003) and yeast (Eide et al. 2005). Similar to the way a photograph depicts two-dimensions of a three-dimensional world, a PCA Biplot is likewise a two dimensional depiction of an N^{th} dimensional world. As such, the relationships defined between variables are constrained by the 'perspective' of the principle components. Nevertheless it is an important entry point to begin understanding multivariate relationships that exist in a correlated phenotypic network like the ionome. Chen et al. (2009) similarly assayed multivariate relationships between ionomic phenotypes using natural variants of *Lotus japonicas*, and wisely pointed out that, unlike in a mutagenized population of what was once a single genotype, more than one unrelated pathways may co-exist to govern elemental accumulation and distribution in a broad survey of natural variants. Considering that

wild proto-*Indica* and proto-*Japonica* forms of *O. rufipogon* diverged more than 100,000 years ago (Ma and Bennetzen 2004) and were independently cultivated by different people in different geographic regions of the world (Fuller et al. 2007, Fuller et al. 2009, Fuller 2011, Fuller et al. 2010, Gu et al. 2013, Hosoya et al. 2010, Sagart 2011, Sweeney and McCouch 2007), it is not unreasonable to assume that different strategies may have evolved convergently in the two groups for dealing with similar stresses. In fact, genetic precedent for this in rice has been established previously relative to aluminum toxicity tolerance where it was shown that different Al tolerance strategies can be employed across subpopulations which has potential importance for breeding (Famoso et al 2011).

Roots are the front line of nutrient acquisition, and are usually where the plant's partitioning decisions are first made. It follows that the correlations between ionomic phenotypes in roots would tend to be more pronounced than in the shoots. For roots (Figure 3.4), there is a clear bi-directional split along PC1 that explains a very large proportion of the variation in the data (69.9%). It also shows a very strong positive correlation between many of the mineral elements in the analysis (cluster II). Revealingly, the dry-weight of the root system (symbolized DW) was included in the analysis and is in strong opposition to cluster II. As part of the analysis pipeline, entire root systems were analyzed rather than sub-sampled in order to maximize the spectral signal, especially from the trace elements, when run through the ICP optical emission spectrometer. As such, in order to derive the concentration of each element in the sample expressed in g/kg (ppm), the undiluted sample values are divided by the

weight of the tissue in the analysis. Despite including root dry weight biomass in the Bayesian modeling of phenotypic means, it is our belief that including the dry weight of the root system in the denominator of the biologically interpretable metric 'parts per million' (ppm) induces a negative correlation between every mineral phenotype in the analysis and the weight of the root system. In turn, this also creates an artifactual, positive correlation between the ionomic phenotypes because they are often negatively correlated with a single variable (Pearson 1896). As such, separation of variables along PC2 in roots might be more informative in terms of defining biologically relevant relationships among members of the ionome.

That being the case, it is interesting to note a few elements that do not exhibit this behavior. The elements in cluster I (namely Mo, B, Rb, Ni, and Li) all share a reasonably parallel unidirectional relationship with root dry weight, while Cu (and to an extent, Pb and Se) share a reasonably perpendicular relationship with root dry weight. Rb, Ni, and Li are among the elements in the analysis with less precise phenotypic measurements due to the analytical limitations of our platform and they tend to exhibit unpredictable relationships to other variables (phenotypic and genetic). Relationships with these phenotypes are likely stochastic, due to the quantity of missing data points that comprise them. Mo and B, on the other hand, are measured with high accuracy and precision and their positive correlation with dry weight may be a relationship with biological underpinnings.

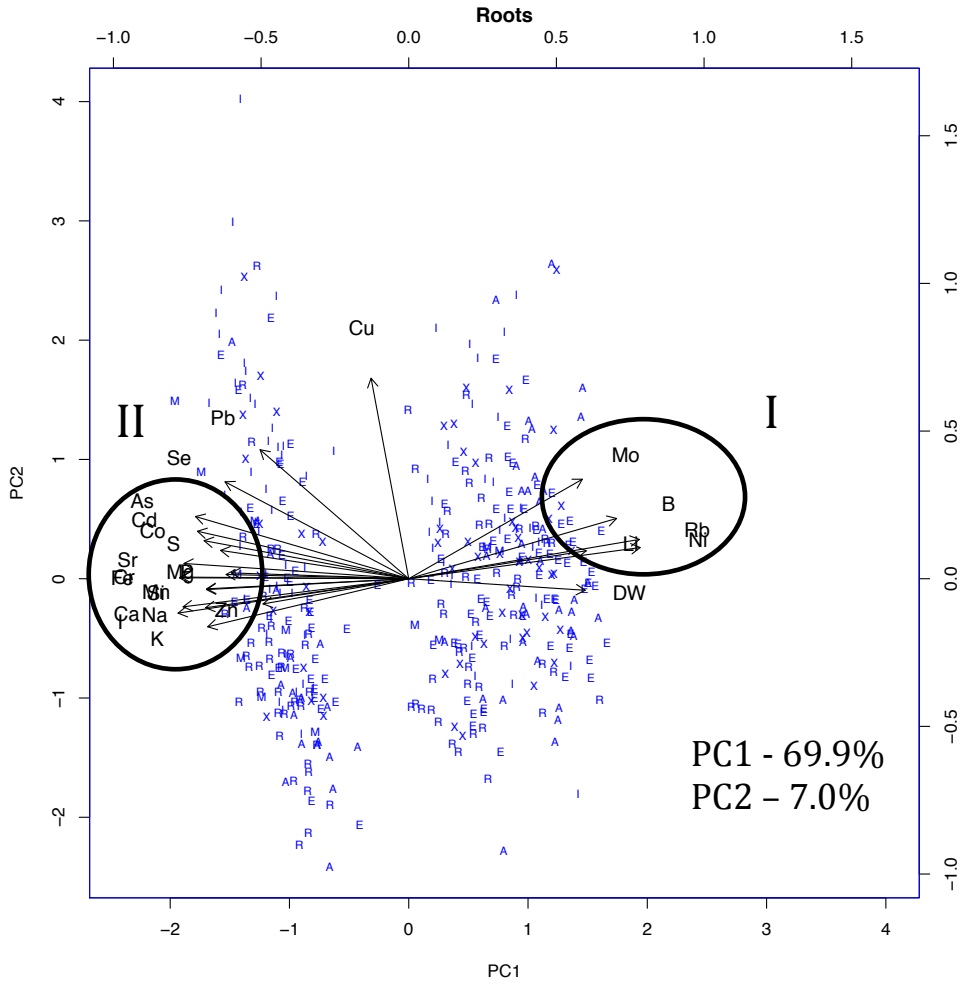


Figure 3.4 PCA Biplot of the rice root ionome. Eigenvectors are displayed for each phenotype in the space defined by the first two principle components. The sub-population identity is given for each of the points representing different accessions (I = *indica;* A = *aus,* R = *tropical japonica*, E = *temperate japonica*, M = *aromatic (Group V)*, and X = admixed lines).

For shoots (Figure 3.5), the elemental analysis was conducted on a subsample of the entire dried homogenized shoot. As such, the variance in the weight of the tissue that was analyzed was low and as none of the relationships depicted are caveated by an induced correlation. In fact, the correlation between whole shoot dry weight biomass and ionomic concentration ranges from a varying negative correlation with a number of elements (cluster III) to no correlation with other phenotypes (clusters I and II). As such, PC1 and PC2 explain a much lower percentage of the variation in shoot tissue (19% and 13.4% respectively). Every principle component after PC2 explains less than 10% of the variation, but twelve principle components are needed before 80% of the variation is described. A scree plot analysis indicates that approximately 5 principle components are needed before the variance explained becomes so low enough to disregard. The fact that 5 dimensions of 25^{th} dimensional space are needed to describe most of the phenotypic variation of the ionome in the diversity panel highlights the complexity and depth of the relationships shared among traits in this correlated phenotypic network. It is also an indication that it will be likely challenging to define predictable physiological trade-offs between elements within the network without considering the rest of the ionome because the relationship between any two elements is highly dependent on other ionomic phenotypes.

On a comparative note, it is interesting that in root tissue, Mo and B are positively correlated, but once partitioning to the shoot has taken place, they are negatively correlated in the shoots.

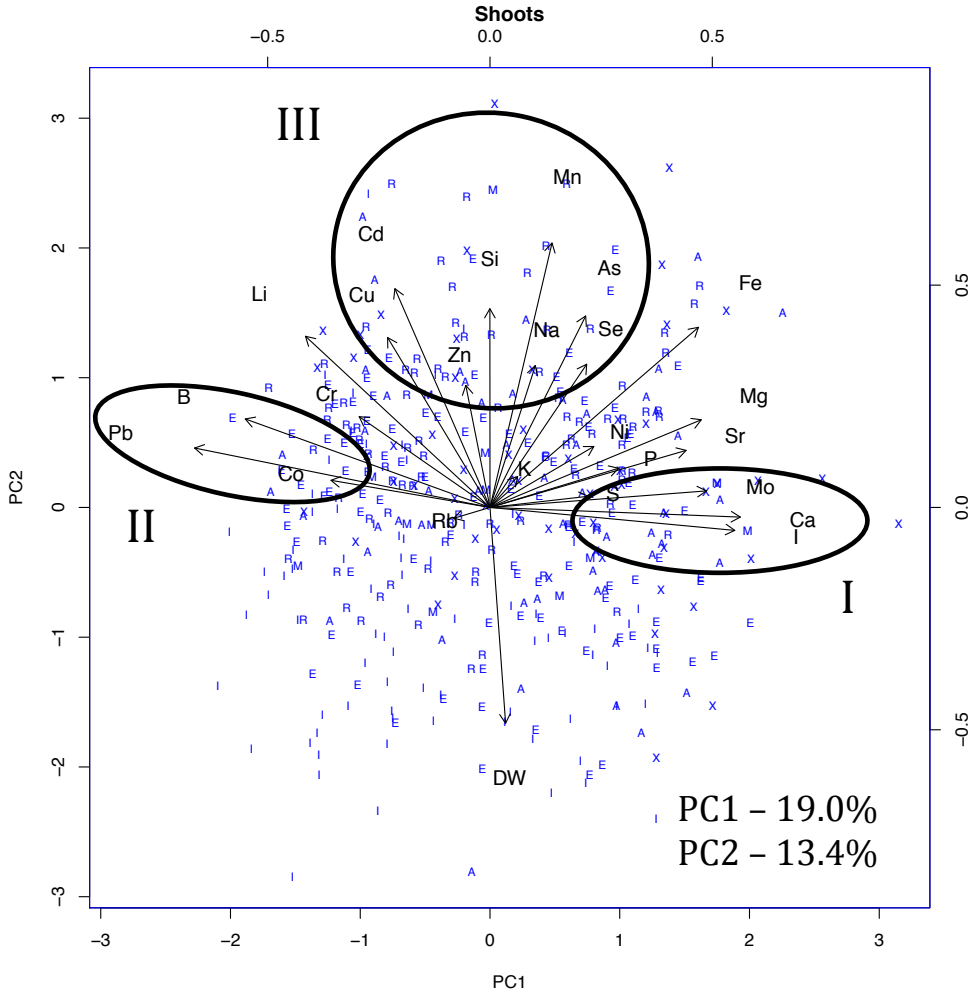


Figure 3.5 PCA Biplot of the rice shoot ionome. Eigenvectors are displayed for each phenotype in the space defined by the fist two principle components. The sub-population identity is given for each of the points representing different accessions (I = *indica;* A = *aus,* R = *tropical japonica*, E = *temperate japonica*, M = *aromatic (Group V)*, and X = admixed lines).

Genome-wide association analysis

Understanding relationships among traits in a correlated network of phenotypes is an important hypothesis generating activity that helps to further the understanding of physiological networks and predict elemental trade-offs, but it does little to expose the underlying genetic basis for the observed relationships. To this end, genome-wide association (GWA) analysis for the ionome was conducted on the entire rice diversity panel as a whole, and individually on each of the corresponding subpopulations. Recently, a custom-designed high density Affymetrix fixed array with over 1,000,000 features was used to genotype the 394 lines in this study (S. McCouch, Cornell University, personal communication). Efficient Mixed Model Association eXpedited (EMMAX; (Kang et al. 2010)) was used to control for population structure and associate SNP genotypes (with minor allele frequencies (MAF) greater than five percent) to each of the univariate rice root and shoot ionomic phenotypes (See Supplementary Figures 3.3 – 3.102 for Manhattan plots, Quantile-Quantile plots, and phenotype distributions associated with the GWA analysis). A 'peak' in the GWA analysis was defined as having at least 3 SNPs within less than 500 kb of one another where, each SNP has a p-value of less than $1x10^{-4}$ and was greater than 500 kb from the next/previous peak. Using these critieria, a total of 362 significant regions were defined across the rice genome for all 50 phenotypes (24 elements plus dry weight biomass in roots and shoots; Supplemental Table 3.1). The average width of a peak was about 364kb and greater than 95% of the peaks were less than 1 MB in width (Figure 3.6). This is approaching the range of reported linkage disequilibrium for domesticated rice which has been calculated to be between 50-500kb (Garris et al.

2005, Mather et al. 2007, Olsen et al. 2006, Rakshit et al. 2007). Table 3.3 shows there were 24 regions that were coincident (within 500 kb of each other) across subpopulations for 16/48 elemental phenotypes. Nineteen of these 24 regions represent coincident regions between the entire population (ALL) and one or more subpopulations. Four of the 24 regions (all for K shoot concentration) were coincident only among subgroups in the *Indica* clade (*indica* and *aus)*, and one curious region for Fe root concentration is found to be coincident between *temperate japonica* and *aus*, possibly representing an *Indica/Japonica* introgression at that region.

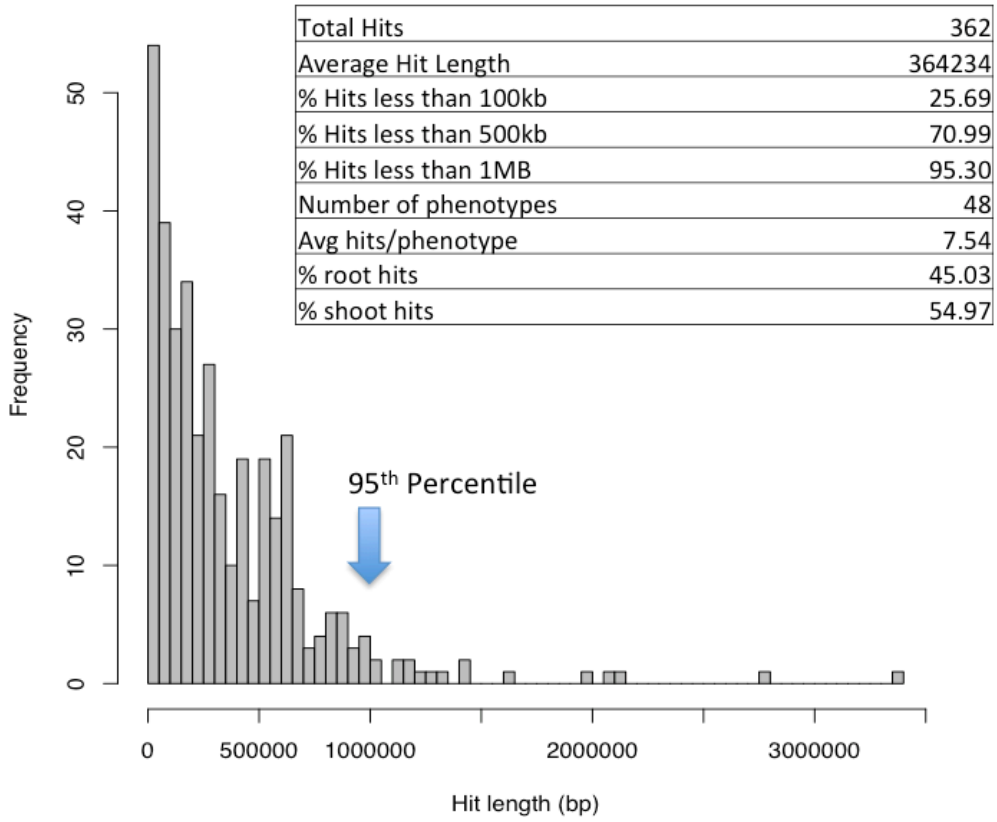


Figure 3.6 Histogram of peak length for GWA analysis and summary statistics.

Identifier	Phenotype	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP Position	Number of Significant SNPs in the region
1	Cd-shoot	ALL	6	29188206	29500577	5.02	29416846	4
	Cd-shoot	TRJ	6	29588432	29642473	4.58	29588432	3
2	Fe-shoot	IND	6	26886515	27106735	4.79	26886515	5
	Fe-shoot	ALL	6	27244620	27365257	6.59	27250341	7
	Fe-shoot	AUS	6	27245820	27365257	4.93	27250341	4
3	Fe-root	TEJ	8	408438	1010352	5.30	885465	9
	Fe-root	AUS	8	636916	637101	5.62	637006	3
4	l-root	AUS	6	27209060	27366961	5.58	27365257	14
	l-root	ALL	6	27244620	27346700	4.76	27250341	4
5	K-shoot	IND	4	58864	458929	4.08	58864	3
	K-shoot	AUS	4	154487	590469	5.43	278717	6
6	K-shoot	AUS	6	7562663	8017951	4.78	7562663	4
	K-shoot	IND	6	7692740	7760675	4.49	7760675	3

Table 3.3 Summary of GWA regions coincident between subpopulation analyses. All positions are given in basepairs (bp) and are defined using the MSU7 assembly of the rice genome.

Identifier	Phenotype	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP Position	Number of Significant SNPs in the region
7	K-shoot	AUS	1	27995687	30115859	4.97	29463693	7
	K-shoot	IND	1	29463693	30339697	6.77	29463693	7
8	K-shoot	AUS	11	20379701	21411524	5.40	20481479	5
	K-shoot	IND	11	21172654	21323271	6.68	21323271	3
9	K-root	ALL	6	888848	1290375	4.30	1290375	3
	K-root	AUS	6	933835	1344591	4.74	1271352	5
10	K-root	AUS	6	27181412	27366961	5.51	27365257	40
	K-root	ALL	6	27240812	27347992	5.29	27250341	11
11	Li-root	AUS	6	676614	1212717	4.78	933835	18
	Li-root	ALL	6	888758	1290375	5.91	1290375	12
12	Mn-root	ALL	1	28988283	30388745	5.88	29758550	60
	Mn-root	TRJ	1	29007814	29816172	5.09	29758550	6

Table 3.3 (continued)

Table 3.3 (continued)

Identifier	Phenotype	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP Position	Number of Significant SNPs in the region
13	Mn-root	ALL	3	28397341	29088366	6.65	29079679	7
	Mn-root	TRJ	3	28653856	29194977	5.49	29194977	6
14	Mn-root	TRJ	7	19193423	19709161	6.68	19700691	27
	Mn-root	ALL	7	19206302	19700691	4.52	19700691	4
15	Mn-root	TRJ	12	24513286	25256090	7.93	24777604	31
	Mn-root	ALL	12	24513286	24874445	7.62	24590841	7
16	Mn-shoot	AUS	10	1632107	2086124	5.66	1812624	5
	Mn-shoot	ALL	10	1925788	2019805	6.85	1965441	10
17	Mo-shoot	IND	8	9952	118947	4.64	116154	10
	Mo-shoot	ALL	8	15243	648977	15.39	200233	152
	Mo-shoot	TRJ	8	16922	320915	6.37	200233	69
18	Na-shoot	ALL	8	7451726	7518485	4.93	7488138	12
	Na-shoot	TEJ	8	7737265	8305931	4.23	7941068	3
	Na-shoot	AUS	8	6340786	7463918	5.76	7132839	7

Table 3.3 (continued)

Identifier	Phenotype	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP Position	Number of Significant SNPs in the region
19	Na-shoot	AUS	8	27468044	28059167	8.41	27585649	28
	Na-shoot	ALL	8	27470977	28060458	7.55	27504822	27
20	Ni-root	AUS	6	27209060	27366961	5.17	27365257	14
	Ni-root	ALL	6	27244620	27346700	4.78	27250341	4
21	P-shoot	TEJ	1	1874689	2061535	6.63	2001792	79
	P-shoot	ALL	1	2001792	2002635	5.12	2002080	3
22	Pb-root	IND	5	7873620	8332608	4.66	7914363	3
	Pb-root	ALL	5	7873620	7942316	5.96	7880314	4
23	Rb-root	AUS	6	27209060	27365593	5.10	27365257	10
	Rb-root	ALL	6	27244620	27346700	4.77	27250341	4
24	S-shoot	ALL	5	19956751	20436357	4.19	20436357	3
	S-shoot	TEJ	5	20434408	20439744	5.66	20434408	4

Likewise, seven regions were identified as being coincident between roots and shoots for the same ionomic phenotype (Table 3.4). If validated, genes in these regions could be regulator genes that govern physiological pathways that control uptake of elements from the soil into the root, as well as other cellular cascades that repartition that element to the shoot. There were also 11 regions identified where five or more peaks overlapped for multiple phenotypes and were designated as genomic hotspots for ionomic phenotypes (Table 3.5). None of these regions were exclusive to phenotypes that would be predicted *a priori* based on elemental species that are chemical analogs of one another (i.e., micronutrients and heavy metals). Eight of the eleven hotspots are for root phenotypes (hotspots 1, 2, 3, 5, 6, 7, 9, and 10), which could suggest that the associations between genotypes at these loci are driven by the strong correlation with root biomass exhibited by most of the root phenotypes. However, only hotspot 8 (which was not enriched for root phenotypes) is also coincident with a GWA region significant for the root biomass phenotype, and even so the peak for root biomass in this region was only detected in the *temperate japonica* subpopulation, while all of the peaks that exhibit association in hotspot 8 are associated with their respective phenotypes only when the entire population is analyzed (see Supplementary Table 3.2 for a complete summary of peaks related to biomass). Additionally, five of the eleven hotspots (1, 3, 5, 6, and 10) are significantly enriched for peaks that were detected only in one subpopulation. Considering that master regulator genes that govern downstream activation of other loci in response to ionic stimulus have been known to vary substantially across subpopulations (Famoso et al. 2011), this further supports the

hypothesis that these hotspots, if validated, could indicate regions of the rice genome that have broad and significant impact on the size, breadth, and distribution of the rice ionome.

Prior to this study the bi-parental QTL analysis offered by (Norton et al. 2010) provided the most comprehensive attempt to describe the underlying genetic architecture of the rice ionome (17 elements). In their analysis they screened both the leaf and grain ionome of 105 F_6 recombinant inbred line (RIL) plants derived from a cross between two divergent varieties (*indica/aus*: cv. Bala and a traditional upland *tropical japonica*: cv. Azucena) genotyped with 164 markers and grown in a flooded paddy at Huazhong Agricultural University in Wuhan, China. (It's important to note here that the *indica/aus* designation of Bala is due to Bala being classified as an *indica* but the particular line used as a parent in this population is much more *aus*-like than other Bala genotypes). Their analysis yielded 36 QTL for leaves and 41 QTL for grain. Based on the mapped position of the peak markers in their analysis, and assuming one centi-Morgan in rice is equal to 240,000 base pairs, nine of their leaf QTL and seven of their grain QTL co-localized with 21 different GWA peaks for the same phenotypes in our analysis (Table 3.6). Most notably, peak 164 (from Supplementary Table 3.1) for As shoot concentration in the *aus* subpopulation is less than 400 kb from the approximated position of QTL qAs5 from Norton et al. (2010) for leaf As content. Interestingly, of the 8 GWA peaks in our analysis discovered for root content of Cu, Fe, Mg, and Ni that co-localized with QTL from Norton et al. (2010), six of them were coincident with QTL for grain concentration and the

remaining two were coincident only with putative leaf QTL (i.e. QTL crossing a threshold based on 1,000 permutations at alpha=0.1 as opposed to the standard alpha=0.05). Also of note is that of the 13 GWA peaks in our analysis for shoot As, Cd, K, Mg, Ni, and Se content that aligned with QTL for their respective phenotypes, four of them (two Cd peaks, one Mg and one Se peak) aligned with grain QTL while the remaining nine peaks aligned with leaf QTL. This again suggests the important implication that the first step to grain accumulation is remobilization from the roots to the shoots, and these QTL may be indicators of the portion of the genetic variance that controls both shoot and grain ionic concentration.

Understanding the alignment of significant loci in our study with the previously published QTL analysis is not only a critical step toward validating the genome-wide association analysis, but because there was such high co-localization between shoot GWA hits and grain QTL the comparison also lends itself to a greater understanding of the partitioning of ionomic constituents to the edible tissues in rice, which is of paramount importance to breeding efforts, especially as those efforts intersect with human health and nutrition (as in the case of Cd and Se above). The comparison of the present study with (Norton et al. 2010) not only lends believability to the GWA peaks discovered here, but also implies that the genetic determinacy for the accumulation of elements of critical importance to human health, such as Cd and As, is tractable and may even be predictable based on the phenotyping of young, non-edible plant tissues.

Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP Position	Number of significant SNPs in region
1	As	root	TEJ	1	23315779	23643681	5.616210392	23366321	14
	As	shoot	IND	1	23395262	23546960	4.754157195	23398195	6
2	Fe	root	ALL	7	22261483	22271232	4.937540224	22263900	3
	Fe	shoot	TEJ	7	22510121	23178153	4.951078169	22515424	7
3	I	shoot	AUS	8	16922	85414	4.362320961	16922	9
	I	root	TEJ	8	408438	1010352	5.659994311	885465	5
4	K	shoot	AUS	6	26738864	26883500	4.212778568	26738864	3
	K	root	AUS	6	27181412	27366961	5.507935531	27365257	40
	K	root	ALL	6	27240812	27347992	5.290313456	27250341	11
5	Li	shoot	IND	2	4561293	4572528	4.328831661	4561293	3
	Li	root	IND	3	4732523	4751762	4.328142056	4732523	3
6	Pb	root	IND	5	25725424	25768128	4.669561269	25768128	9
	Pb	shoot	ALL	6	25887619	26063255	4.390428787	25887619	3
7	Zn	shoot	AUS	8	19282078	19575921	5.186467572	19575921	19
	Zn	root	TEJ	11	19233267	19425358	6.380354744	19425358	3

Table 3.4 Summary of GWA peaks coincident across tissues. All positions are given in basepairs (bp) using the MSU7 rice genome assembly.

Hotspot Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP Postion	Number of significant SNPs in region
1	As	root	TEJ	1	23315779	23643681	5.616210392	23366321	14
	Ca	root	TEJ	1	23346777	23591173	5.105155675	23591173	8
	Cd	root	TEJ	1	23338971	23643681	5.923997604	23523041	15
	Cr	root	TEJ	1	23338971	23624147	6.007062298	23591173	15
	Fe	root	TEJ	1	23343063	23624147	5.603130083	23591173	12
	Mn	root	TEJ	1	23315779	23643681	8.041833118	23591173	21
	Na	root	TEJ	1	23346777	23591173	4.635126498	23523041	4
	Ni	root	TEJ	1	23366321	23591173	4.901362442	23591173	4
	Sr	root	TEJ	1	23341359	23643681	5.826162925	23501367	15
	As	shoot	IND	1	23395262	23546960	4.754157195	23398195	6
	I	shoot	TEJ	1	22980639	23084851	4.372753147	22987492	3
	Mg	shoot	IND	1	22640125	22913650	4.993243218	22640125	8
	Pb	shoot	TEJ	1	23338971	23523041	4.648011543	23343063	9
2	Co	root	ALL	1	30211322	30290566	4.558659611	30211322	3
	I	root	ALL	1	29145484	29183500	4.198595016	29183500	3
	Mn	root	ALL	1	28988283	30388745	5.884480132	29758550	60
	Mn	root	TRJ	1	29007814	29816172	5.090428038	29758550	6
	Ni	root	ALL	1	29145484	29183500	4.235000261	29183500	3
	P	root	IND	1	29110936	29120311	4.424297196	29118033	3
	Pb	root	IND	1	29050028	29150350	4.920031829	29118033	9
	Rb	root	ALL	1	29122296	29499909	4.566466272	29183500	8
	Se	root	AUS	1	29615659	30130607	4.829511982	29947864	9
	Sr	root	IND	1	28924884	29303536	5.246232286	29122296	7
	K	shoot	IND	1	29463693	30339697	6.771358816	29463693	7

Table 3.5 Summary of GWA hotspots. All positions are given in basepairs (bp) using the MSU7 rice genome assembly.

Hotspot Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP	Number of significant SNPs in region
3	B	root	IND	3	30874384	31303675	5.058417127	31179006	68
	Ca	root	IND	3	31126028	31303675	5.415887935	31298364	54
	Cr	root	IND	3	31179006	31298364	5.214186877	31265273	17
	Fe	root	IND	3	31249805	31298364	4.248185263	31249805	6
	I	root	IND	3	31179006	31298364	4.838044505	31265273	13
	Li	root	IND	3	31113733	31303675	5.370701383	31265273	22
	Na	root	IND	3	31170812	31303675	5.337558795	31265273	17
	Ni	root	IND	3	31265273	31298364	4.634803633	31265273	5
	Rb	root	IND	3	31179006	31298364	4.903274509	31265273	17
	Sr	root	IND	3	31126028	31294481	5.046209942	31210807	37
	K	shoot	AUS	3	30916945	31070960	4.960516506	30916945	3
4	Li	root	ALL	5	24906165	25057270	4.559104719	24973819	5
	Li	root	ALL	5	25874780	26980454	7.29431423	26250170	42
	Pb	root	IND	5	25725424	25768128	4.669561269	25768128	9
	Zn	root	TEJ	5	24346854	24583903	6.55933594	24471818	5
	K	shoot	AUS	5	25146331	25700676	4.991801139	25700676	4
	Ni	shoot	ALL	5	24973819	25206799	6.183898964	25054618	10
	Sr	shoot	ALL	5	25839893	25939552	4.873609786	25839893	3

Table 3.5 (continued)

Hotspot Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP	Number of significant SNPs in region
5	Ca	root	AUS	6	665738	1290375	6.195127067	1125502	76
	Cr	root	AUS	6	665738	1344591	6.467518241	1069445	78
	Fe	root	AUS	6	665738	1290375	5.998486101	1125502	53
	I	root	AUS	6	802340	1344591	5.5824812	933835	41
	K	root	ALL	6	888848	1290375	4.303045609	1290375	3
	K	root	AUS	6	933835	1344591	4.737454029	1271352	5
	Li	root	AUS	6	676614	1212717	4.784971618	933835	18
	Li	root	ALL	6	888758	1290375	5.908328827	1290375	12
	Mo	root	ALL	6	888848	1290375	5.781872289	1290375	6
	Ni	root	AUS	6	802340	1344591	5.376469806	1125502	38
	P	root	AUS	6	802340	838315	4.48534069	819994	3
	Rb	root	AUS	6	802340	1344591	5.196496688	933835	37
	S	root	AUS	6	948665	1290375	5.495614441	1222551	19
	Si	root	AUS	6	1048700	1069445	4.063985087	1058312	3
	Sr	root	AUS	6	802340	894443	4.58904152	802340	8
	Se	shoot	IND	6	725500	744876	5.664832204	742145	6

Table 3.5 (continued)

Hotspot Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP	Number of significant SNPs in region
6	B	root	AUS	6	27297876	27591816	5.806720837	27365257	12
	Cr	root	AUS	6	27070507	27365257	4.786883296	27365257	6
	I	root	AUS	6	27209060	27366961	5.58447615	27365257	14
	I	root	ALL	6	27244620	27346700	4.755840754	27250341	4
	K	root	AUS	6	27181412	27366961	5.507935531	27365257	40
	K	root	ALL	6	27240812	27347992	5.290313456	27250341	11
	Li	root	AUS	6	27209060	27365257	4.848123639	27209060	5
	Ni	root	AUS	6	27209060	27366961	5.170930206	27365257	14
	Ni	root	ALL	6	27244620	27346700	4.777613842	27250341	4
	Rb	root	AUS	6	27209060	27365593	5.098431888	27365257	10
	Rb	root	ALL	6	27244620	27346700	4.76776274	27250341	4
	Fe	shoot	IND	6	26886515	27106735	4.789448678	26886515	5
	Fe	shoot	ALL	6	27244620	27365257	6.593552159	27250341	7
	Fe	shoot	AUS	6	27245820	27365257	4.926699294	27250341	4
	K	shoot	AUS	6	26738864	26883500	4.212778568	26738864	3
7	K	root	TRJ	7	18658838	19992804	4.674948609	19991889	14
	Mn	root	TRJ	7	19193423	19709161	6.679731604	19700691	27
	Mn	root	ALL	7	19206302	19700691	4.516035563	19700691	4
	P	root	ALL	7	19554097	19557934	5.386205334	19556364	3
	I	shoot	ALL	7	18508120	18654861	5.402764319	18654861	4
	Ni	shoot	ALL	7	19958320	20561696	5.587249907	20084671	8

Table 3.5 (continued)

Hotspot Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP	Number of significant SNPs in region
8	B	root	ALL	7	22826217	22924664	4.93008682	22897553	6
	Cr	root	ALL	7	22263900	22296846	4.648734627	22263900	3
	Cr	root	ALL	7	22894666	22924664	4.815244757	22897553	4
	Fe	root	ALL	7	22261483	22271232	4.937540224	22263900	3
	Cd	shoot	AUS	7	21178877	21992234	6.540870653	21178877	5
	Fe	shoot	TEJ	7	22510121	23178153	4.951078169	22515424	7
	I	shoot	TEJ	7	22515424	22924664	4.771796177	22916565	5
	Ni	shoot	IND	7	22680383	23231066	5.957412543	23229073	4
9	Cu	root	AUS	7	28664814	28667427	5.11609604	28666781	5
	P	root	IND	7	28146025	28750621	4.504493808	28415622	8
	Pb	root	IND	7	29142588	29420306	4.578085308	29142588	3
	S	root	AUS	7	29409744	29414791	4.415074589	29414791	3
	Co	shoot	TRJ	7	28706880	29248197	4.445451839	29248197	4

Table 3.5 (continued)

Hotspot Identifier	Phenotype	Tissue	Sub-pop	Chr	Start (bp)	Stop (bp)	Max Negative LOG(p-value)	Peak SNP	Number of significant SNPs in region
10	Ca	root	TEJ	8	408438	1010352	7.134557805	885465	13
	Cr	root	AUS	8	636916	637101	6.146407967	637006	4
	Fe	root	TEJ	8	408438	1010352	5.299578544	885465	9
	Fe	root	AUS	8	636916	637101	5.624420575	637006	3
	I	root	TEJ	8	408438	1010352	5.659994311	885465	5
	Li	root	TEJ	8	408438	1010352	4.521254537	885465	3
	Ni	root	TEJ	8	371458	1010352	5.474094766	885465	6
	P	root	TEJ	8	408438	878120	4.569671065	878120	3
	Rb	root	TEJ	8	408438	1010352	5.539548986	885465	5
	Sr	root	TEJ	8	408438	935991	4.543128143	885465	3
	I	shoot	AUS	8	16922	85414	4.362320961	16922	9
	K	shoot	IND	8	527192	560290	4.125554583	555392	6
	Mo	shoot	IND	8	9952	118947	4.643899383	116154	10
	Mo	shoot	ALL	8	15243	648977	15.39310871	200233	152
	Mo	shoot	TRJ	8	16922	320915	6.367154702	200233	69
11	Zn	root	IND	8	6708187	7320736	4.802501089	7161287	3
	Mg	shoot	IND	8	8772106	8805623	5.04347668	8772106	7
	Na	shoot	AUS	8	6340786	7463918	5.763440163	7132839	7
	Na	shoot	ALL	8	7451726	7518485	4.928597518	7488138	12
	Na	shoot	TEJ	8	7737265	8305931	4.23134726	7941068	3

Table 3.5 (continued)

QTL from Norton et al. 2009	Peak Identifier	Phenotype	Sub-pop	Chr	Peak Start	Peak End	Approximate distance of QTL from Peak Start (MB)
qAs3-LEAF	166	As-shoot	IND	3	35410744	35589970	-2.284
qAs5-LEAF	164	As-shoot	AUS	5	21268077	21554491	0.647
qCd7-GRAIN	179	Cd-shoot	AUS	7	424382	545121	5.501
qCd9*-GRAIN	173	Cd-shoot	ALL	9	14814169	15013054	5.843
qCu2-GRAIN	29	Cu-root	TRJ	2	24378805	24459501	-3.048
	22	Cu-root	AUS	2	25761069	25827209	-4.431
qFe3-GRAIN	33	Fe-root	IND	3	31249805	31298364	-2.203
qFe6*-LEAF	31	Fe-root	AUS	6	665738	1290375	1.896
qK9.1-LEAF	252	K-shoot	AUS	9	9605468	9993344	-3.348
qK9.2-LEAF	253	K-shoot	AUS	9	17114488	18143973	-1.791
qMg1-LEAF	285	Mg-shoot	AUS	1	35343330	35494441	1.171
qMg3-GRAIN	287	Mg-shoot	IND	3	2273913	2298354	-1.156
	67	Mg-root	AUS	3	3759765	3768164	-2.642
qMg6-LEAF	289	Mg-shoot	TEJ	6	3562814	3602700	-1.740
qNi1.2*-LEAF	326	Ni-shoot	AUS	1	18690404	18725405	4.082
	106	Ni-root	TEJ	1	23366321	23591173	-0.594
qNi7*-LEAF	325	Ni-shoot	ALL	7	19958320	20561696	-0.572
	328	Ni-shoot	IND	7	22680383	23231066	-3.294
qNi9*-GRAIN	103	Ni-root	AUS	9	444217	1399643	4.853
	104	Ni-root	AUS	9	1904206	2500461	3.393
qSe1.2-GRAIN	348	Se-shoot	ALL	1	41244480	41587329	1.409

Table 3.6 Comparison of GWA Peaks with QTL from Norton et al. (2010). All positions are given using the MSU7 assembly of the rice genome. A minus symbol in the approximate position indicates the QTL is upstream from the coincident peak. * indicates QTL significant only at alpha=0.05.

Co-localization of GWA peaks with *a priori* candidate genes

To begin to explore and annotate the GWA mapping results, a list of rice candidate genes for each phenotype was compiled using gene ontology (GO) terms related to the transport and movement of ions across plant cell membranes and throughout the plant. The same GO terms were also used to probe the *Arabidopsis thaliana* and *Zea mays* genomes for putative candidates. Rice orthologs of those genes were then identified and classified as putative candidates using the Ensembl COMPARA pipeline (www.ensembl.org/info/docs/compara/index.html). Combining the candidates identified by searching the rice genome directly, along with the putative candidate genes identified by orthology with *A. thaliana* and *Z. mays,* a total of 1,222 unique rice gene models were implicated for all of the ionomics phenotypes.

The genes identified here provide an important, if preliminary and necessarily incomplete, genetic infrastructure upon which meaningful interpretation of GWA peaks can be conducted. This also provides a mechanism to quickly identify peaks for follow up analysis and targeted, hypothesis driven validation. In addition to the putative candidate genes, a list of 28 candidate genes, cloned and characterized in rice was also compiled (Table 3.7). All of the candidate genes cloned in rice for a given phenotype were combined with putative candidates for the same phenotype that fell within 500kb of a GWA peak and aligned with the distribution of SNPs on the Manhattan plots to improve the efficiency of peak annotation (See Supplementary Figures 3.3-3.52). Five of the 28 cloned candidate genes (YSL2, ZIP8, ZIP6, ZIP4 and ZIP5) co-localized with association hits for root concentration of Cu, Mn, Pb, and Zn,

respectively. The remaining cloned candidate genes were either not highlighted within a peak or the peak associated with them was not significant enough in our population to meet the useful, but arbitrary significance threshold of $1x10^{-4}$.

A total of 163 GWA peaks for various root and shoot phenotypes aligned with 131 putative candidate ion transport genes (Supplementary Table 3.3), some of which harbored significant SNPs in the gene models themselves (such SNPs are highlighted in red in Supplementary Figures 3.3 – 3.52). While some of this co-localization is expected by random chance (given the probability of a stochastically generated peak and the relative number and distribution of putative candidate genes across the genome for a given phenotype), it still offers compelling evidence that in many cases, GWA analysis is able to identify novel loci.

.

Gene Name	Chr	Start	Stop	Candidacy Phenotypes	Reference
ATX1	2	19466343	19480380	Fe	Ganesh et al. 2002
BOR1	12	23215284	23219805	B	Nakagawa et al. 2007
BOR3	1	3877069	3899085	B	Nakagawa et al. 2007
BOR4	5	4600956	4607390	B	Nakagawa et al. 2007
DMAS1	3	7227962	7230747	Fe	Bashir et al. 2006
FRO1	4	21997480	22001824	Fe	Ishmaru et al. 2006
FRO2	4	28993724	28996525	Fe	Ishmaru et al. 2006
IRT1	3	26279204	26285071	Cd, Co, Cu, Fe, Ni, Pb,Zn	Bughio et al. 2002
IRT2	3	26274320	26275225	Cd, Co, Cu, Fe, Ni, Pb,Zn	Ishmaru et al. 2006
Lsi1	2	31260376	31264091	Si	Ma et al. 2006
Lsi2	3	429476	431876	Si	Ma et al. 2006
NAAT1	2	11997090	12002629	Fe	Inoue et al. 2008
NAS1	3	10928201	10929613	Fe	Inoue et al. 2008
NAS2	3	10925187	10926668	Fe	Inoue et al. 2008
NAS3	7	29322102	29323731	Fe	Inoue et al. 2008
OsHKT1;5	1	11457931	11462418	Na	Ren et al. 2005
OsHMA3	7	7404750	7408558	Cd	Miyadate et al. 2011
OsLCT1	6	22565778	22570985	Cd	Uraguchi et al. 2011
Tpc1	1	27905564	27919936	Ca	Kurusu et al. 2004
YSL2*	2	26164518	26169101	Cu, Fe, Zn	Koike et al. 2004
ZIP1	1	42904522	42906430	Cd, Co, Mn, Ni, Pb, Zn	Ramesh et al. 2003
ZIP2	3	17006829	17010137	Cd, Co, Mn, Ni, Pb, Zn	Ishimaru et al. 2005
ZIP3	4	30893089	30895447	Cd, Co, Mn, Ni, Pb, Zn	Ramesh et al. 2003
ZIP4*	8	6266828	6269910	Cd, Co, Mn, Ni, Pb, Zn	Ishimaru et al. 2005
ZIP5*	5	23153854	23156724	Cd, Co, Mn, Ni, Pb, Zn	Lee et al. 2010
ZIP6*	5	3807951	3810783	Cd, Co, Mn, Ni, Pb, Zn	Ishimaru et al. 2005
ZIP7	5	5996577	6000540	Cd, Co, Mn, Ni, Pb, Zn	Yang et al. 2009
ZIP8*	7	7393491	7396720	Cd, Co, Mn, Ni, Pb, Zn	Yang et al. 2009

Table 3.7 Cloned rice candidate genes. Gene names marked with an * indicate that they were identified in the GWA analysis. All positions are given using the current (MSU7) rice reference genome assembly.

Annotation of selected GWA peaks - Molybdenum

The most significant GWA peak in the entire GWA analysis is for molybdenum shoot concentration, surveyed across the entire RDP1 (Peak ID 296 in Supplementary Table 3.1; Figure 3.7). This peak is coincident with similarly significant peaks in *indica* and *tropical japonica*. The peak extends across 633 kilobases of the distal region of the small arm of rice chromosome 8, and the negative LOG of the most significant SNP in the peak is 15.39. There are 99 gene models in the reference Nipponbare genome that fall within the region of this peak, and there are no cloned genes in rice known to specifically impact molybdenum concentration.

One of the gene models (LOC_Osg0801120) under the peak was highlighted by the candidate gene analysis as being associated with the gene ontology term GO:0015098 ("molybdate ion transmembrane transporter activity"). This gene is electronically annotated in the rice reference genome as an expressed putative sulfate transporter (sulfate is a known chemical analog of molybdate). The locus LOC_Os08g01120, (heretofore referred to as *OsMOT1*), is 455 residues in length and has a single exon. CLUSTAL W multiple sequence alignment (Larkin et al. 2007) shows 59% protein sequence identity with a similarly structured orthologous gene in *A. thaliana* recently described as *AtMOT1* (AT2G25680; (Baxter et al. 2008, Tomatsu et al. 2007).

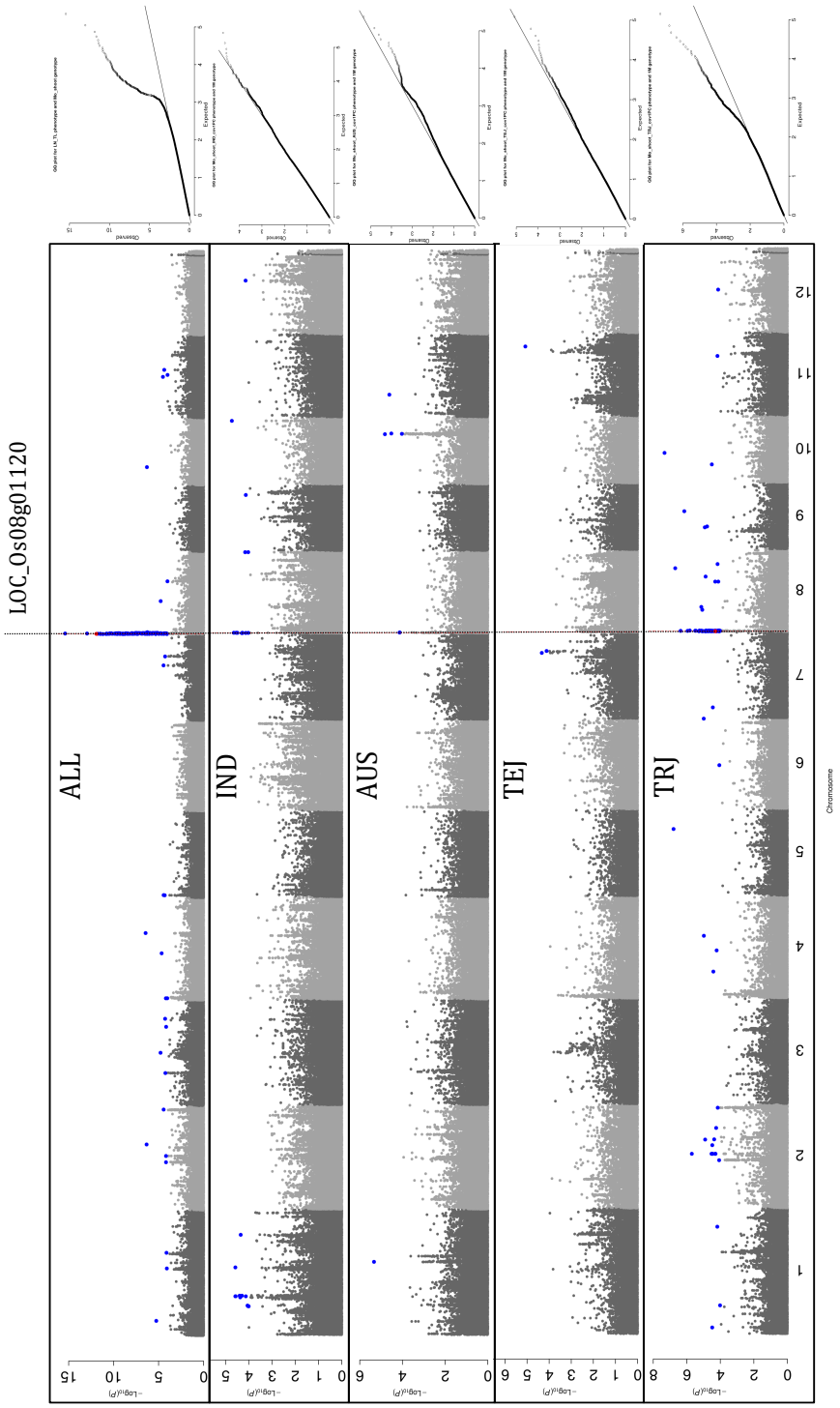


Figure 3.7 Manhattan and quantile-quantile plots for molybdenum shoot content. Blue SNPs have p-values less than 10^{-4}, and red SNPs are in the candidate gene itself. LOC_08g01120 (*OsMOT1*) position is as indicated

Haplotype analysis of *OsMOT1* using the HDRA

Mo shoot concentration in the diversity panel was normally distributed in the RDP1, and 10% of the variation was explained by subpopulation (p-value <0.0001; Figure 3.8). The most notable pattern was observed within the *Indica* clade where the *aus* subpopulation tended to have the highest shoot Mo concentration, and the *indica* subpopulation had the lowest. In an effort to understand the underlying variation at this locus, haplotype analysis across the *OsMOT1* gene was conducted based on the HDRA genotype data for the RDP1 (Figure 3.9). Three SNPs on the array were located within the *OsMOT1* gene model and are found either in the 5' untranslated region (5' UTR), the exon, or the 3' UTR. Furthermore, there were two SNPs on the array in the inter-genic space upstream of *OsMOT1* . Using these 5 markers, nine distinct haplotypes were identified in the RDP1. Only the first four haplotypes were found at frequencies > 5%, and showed significant enrichment within one subpopulation or another. For example, out of 77 *temperate japonica* varieties, 73 contained the 'reference genotype' (i.e., identical to the Nipponbare genome) at all SNPs under consideration. Likewise 44/46 *indica* lines and 46/69 *tropical japonica* lines harbored a variant allele only at the non-synonymous SNP in the exon.

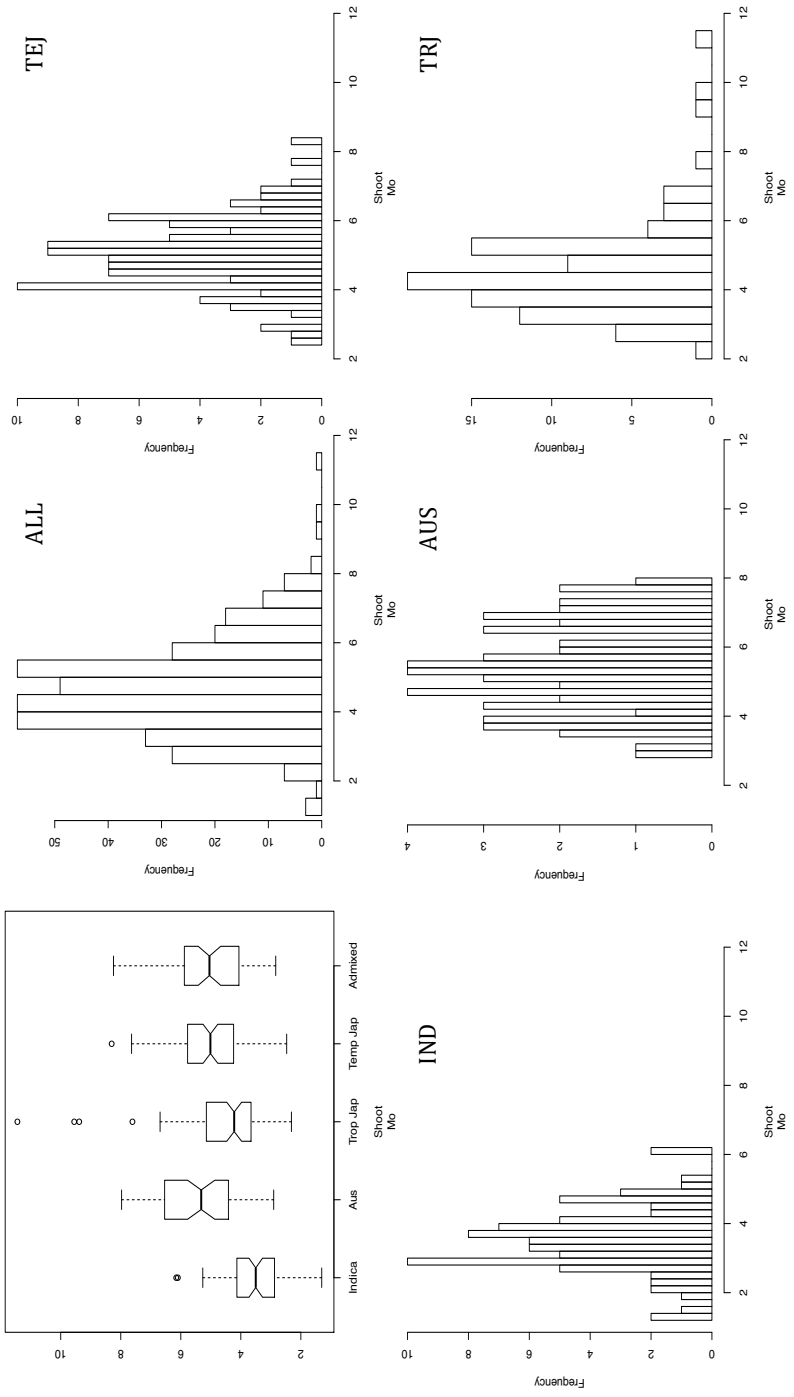


Figure 3.8 Boxplots and histogram distributions of molybdenum shoot phenotype data. Each subpopulation is represented individually, and ALL represents the entire panel. Subpopulation identity explains 9.7% of the phenotypic variation for Mo shoot content (p-value <.00001).

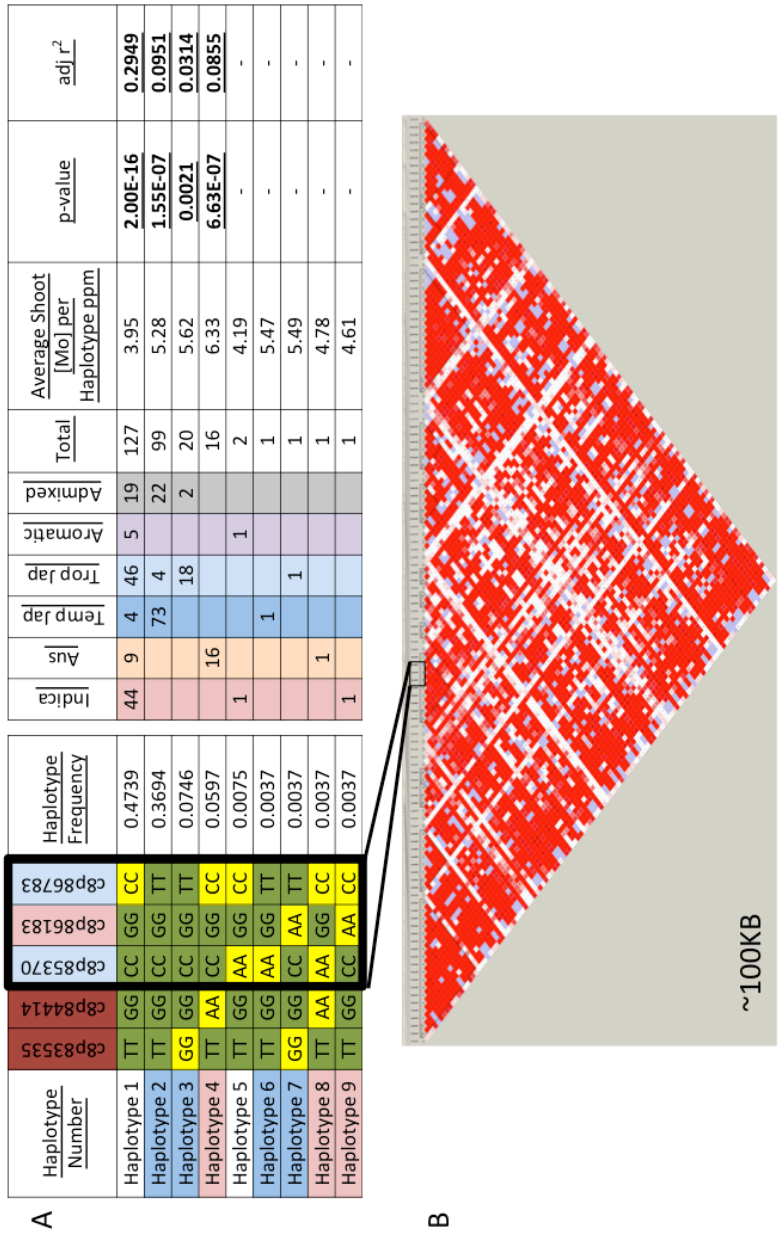


Figure 3.9 LD and Haplotype analysis of region surrounding *OsMOT1* . A) Haplotype analysis of *OsMOT1* using the HDRA. SNPs in the haplotype analysis highlighted in dark red are in the inter-genic region upstream of the gene. SNPs inside the black outline are in the gene model. Light red SNPs are non-synonymous. SNPs highlighted in blue are in the gene model but present in the 3'/5' UTR. Haplotype numbers highlighted in blue are specific to the *Japonica* clade, those in orange are specific to the *Indica* clade. Genotypes in dark green are identical to the reference genome (Nipponbare), Yellow represent the alternate allele. B) LD analysis of 100 kb region surrounding *OsMOT1* based on genotype data from the HDRA on the diversity panel. LD Plot spans the region from chr8:75000 – chr8:100,000.

Only two haplotypes ('Haplotype 3' and 'Haplotype 4') were specific to a particular subpopulation (*tropical japonica* and *aus,* respectively), and offered shoot Mo concentration means significantly higher than the most abundant haplotype in the panel ('Haplotype 1'). Haplotype variation, specifically in the *tropical japonica* subpopulation (which has the strongest peak among subpopulation-specific analyses), was bifurcated primarily between 'Haplotype 1' (66.6% frequency in *tropical japonica*) and 'Haplotype 3' (26% frequency in *tropical japonica*, 7.5% in ALL; Figure 3.10 A). These two haplotypes could be distinguished by either a SNP in the 3'UTR of *OsMOT1* (c8p86783) or by an upstream SNP (c8p83535). Individuals carrying 'Haplotype 3' had significantly higher average Mo shoot concentrations than other haplotypes (5.58 ppm compared to 3.86 ppm) in *tropical japonica*.

Despite the fact that no formal peak was identified in the *aus* specific subpopulation analysis, haplotype variation at *OsMOT1* was similarly bifurcated by two major haplotypes ('Haplotype 1' and 'Haplotype 4'). 'Haplotype 1' was the most abundant haplotype in the entire panel (frequency of 47%), but had a frequency of 35% in the *aus* subpopulation. 'Haplotype 4', which had a frequency of 6% in the RDP1 was found in 62% of lines in the *aus* subpopulation. The locus explained 30.6% of the variation for Mo shoot concentration in *aus* (average of 4.77 ppm Mo for 'Haplotype 1' compared to 6.33 ppm Mo for 'Haplotype 4') (Figure 3.10 B). Interestingly, the lines in the *aus* gene pool are monomorphic for the HDRA SNPs in the coding region of *OsMOT1* , but 'Haplotype 1' and 'Haplotype 4' can be readily distinguished by a single SNP (c8p84414) upstream of the 5' UTR.

Resulting from the fact that most of the subpopulation specific variation for Mo shoot concentration can be explained by two SNPs upstream of *OsMOT1* and one SNP in the 3' UTR, it is hypothesized that differences in gene expression of *OsMOT1* are driving this gene's contribution to the overall phenotypic variance for shoot Mo accumulation.

TRJ

A

Haplotype Number	c8p83535	c8p84414	c8p85370	c8p86183	c8p86783	Haplotype Frequency	# TRJ Lines	Average Shoot [Mo] per Haplotype (TRJ Only) ppm	p-value	adj r^2
Haplotype 1	TT	GG	CC	GG	CC	0.6667	46	3.86	**1.88E-07**	**0.3252**
Haplotype 2	TT	GG	CC	GG	TT	0.0580	4	4.98	-	-
Haplotype 3	GG	GG	CC	GG	TT	0.2609	18	5.58	**2.18E-06**	**0.2756**
Haplotype 7	GG	GG	CC	AA	TT	0.0145	1	5.49	-	-

AUS

B

Haplotype Number	c8p83535	c8p84414	c8p85370	c8p86183	c8p86783	Haplotype Frequency	# Aus Lines	Average Shoot [Mo] per Haplotype (*aus* Only) ppm	p-value	adj r^2
Haplotype 1	TT	GG	CC	GG	CC	0.3462	9	4.77	**0.0052**	**0.253**
Haplotype 4	TT	AA	CC	GG	CC	0.6154	16	6.33	**0.002**	**0.3058**
Haplotype 8	TT	AA	AA	GG	CC	0.0385	1	4.78	-	-

Figure 3.10 Subpopulation specific haplotype analysis of *tropical japonica* and *aus*. SNPs in the haplotype analysis highlighted in dark red are in the inter-genic region upstream of *OsMOT1* . SNPs inside the black outline are in the gene model. Light red SNPs are non-synonymous. SNPs highlighted in blue are in the gene model but present in the 3'/5' UTR. Haplotype numbers highlighted in blue are specific to the *Japonica* clade, those in light red are specific to the *Indica* clade. Genotypes in dark green are identical to the reference genome (Nipponbare), Yellow represent the alternate allele.

Sequence Analysis of *OsMOT1* using re-sequenced lines

To identify additional polymorphisms across *OsMOT1* that might refute this hypothesis, full-length sequence variation was compared separately. We compared re-sequencing data (7-50x genome coverage) across the region for 125 rice genomes (McCouch, S and Wright M personal communication), of which 52 lines were also included in the diversity panel, and thus we had both HDRA genotype data and Mo shoot phenotype data for each of them. Sequence variation was compared for the *OsMOT1* coding region and approximately 3 kb of upstream DNA.

For the transcribed portion of *OsMOT1* , there were 3 non-synonymous SNPs in the exon and 5 SNPs in the 3' UTR that were polymorphic among the lines for which both phenotypic data and re-sequencing information was available (Figure 3.11 A). Consistent with the haplotype analysis based on the HDRA data, all of the re-sequenced *temperate japonica* lines were monomorphic for every SNP in the coding region of *OsMOT1* , and the genetic variation at this locus in *tropical japonica* showed only two haplotype groups ('CDS 1' and 'CDS 7'), both of which were found to be specific to the *Japonica* clade and completely monomorphic for the SNPs in the coding region of *OsMOT1*.

More variation was observed among *indica* and *aus* lines, and haplotypes not discovered previously were identified in these groups, including two haplotypes ('CDS 4' and 'CDS 6') in the *indica* subpopulation and one haplotype ('CDS 5') in

the *aus* subpopulation. These haplotypes could be distinguished by non-synonymous SNP variation in the exon of *OsMOT1*. In addition to 'CDS 5', the *aus* subpopulation also exhibited two other haplotypes: 'CDS 3', distinguished by two polymorphisms in the 3' UTR, and 'CDS 2' distinguished by a another single but different SNP in the 3'UTR. Only one haplotype ('CDS 3') was found to be present across clades, and was notably more abundant in the *Indica* varietal group, and specifically the *aus* subpopulation. When all of the lines in the RDP1 were considered together, three haplotypes ('CDS 4', '5', and '7') explained a trivial amount of phenotypic variation.

To investigate the hypothesis that variation in the upstream, promoter region of *OsMOT1* might be driving Mo shoot accumulation, sequence analysis was similarly performed based on 1,908 base pairs upstream of *OsMOT1* (Figure 3.11 B). There were 105 polymorphic SNPs in this "promoter" region for the phenotyped re-sequenced lines, and 28 of them explained a statistically significant amount of the phenotypic variation for Mo shoot content, either in the RDP1 as a whole, and/or when evaluated in a subpopulation-specific manner. These 28 SNPs were subsequently used to create upstream haplotypes. A total of 12 upstream haplotypes were defined using these 28 SNPs. Three haplotypes ('Upstream 1', '3', and '6') were specific to the *Japonica* clade, and seven ('Upstream 5', '7', '8', '9', '10', '11', and '12') were specific to the *Indica* clade. Only two haplotypes were present across varietal groups ('Upstream 2' and '4'), but even then were defined by much higher frequencies in either the *Japonica* clade ('Upstream 2') or the *Indica* clade ('Upstream 4'). Similar to the sequence analysis of the coding region of *OsMOT1* , only 3

upstream haplotypes ('Upstream 5', '6', and '7') explained a significant, but small portion of the phenotypic variance. 'Upstream 7' is an exception, explaining 14.7% of the phenotypic variance, largely due to the fact that it is specific to the *indica* subpopulation and is characteristically found in lines that have low shoot Mo concentrations (~2.8ppm).

A

Haplotype Number	8:86771	8:86946	8:87741+	8:87784+#	8:87981*	8:88233*	8:88288*	8:88316*	Haplotype Frequency	Indica	Aus	Temp Jap	Trop Jap	Aromatic	Admixed	Total	Average Shoot [Mo] per Haplotype	p-value	adj r^2
CDS 1	GG	CC	TT	TT	GG	GG	CC	AA	0.4024			18	12	2	2	33	4.60 [20]	-	-
CDS 2	GG	CC	TT	CC	GG	CC	CC	AA	0.1220	6	2				2	10	4.04 [3]	-	-
CDS 3	GG	CC	TT	CC	GG	GG	GG	AA	0.1098	2	5			1	1	9	4.91 [5]	-	-
CDS 4	TT	CC	TT	CC	GG	GG	CC	TT	0.0976	7					1	8	3.51 [4]	**0.0812**	**0.0485**
CDS 5	GG	AA	TT	CC	GG	GG	CC	AA	0.0976		7				1	8	5.16 [7]	**0.073**	**0.0525**
CDS 6	GG	CC	CC	CC	..	GG	CC	AA	0.0976	5					3	8	3.61 [2]	-	-
CDS 7	GG	CC	TT	CC	AA	CC	CC	AA	0.0732				6			6	3.27 [3]	**0.0608**	**0.0593**

B

Haplotype Number	8:84506**	8:84536**++	8:84570* **	8:84571** **	8:84660*	8:84771*	8:84952**	8:84982**	8:85352**	8:85395** **	8:85399**	8:85415+	8:85444**	8:85450**	8:85465* **	8:85469**	8:85778**	8:85828** **	8:85844**	8:85936* **	8:85937**	8:85964**	8:85967** **	8:86005** **	8:86039**	8:86058**	8:86101* **	8:86126** **	8:86128**
Upstr 1	AA	TT	GG	GG	CC	AA	AA	TT	AA	CC	AA	GG	TT	CC	GG	GG	AA	GG	GG	AA	AA	GG	CC	CC	TT	CC	TT	TT	TT
Upstr 2	AA	TT	GG	GG	CC	GG	AA	TT	AA	CC	AA	GG	TT	CC	GG	GG	AA	GG	GG	GG	AA	GG	CC	CC	TT	CC	TT	TT	TT
Upstr 3	AA	GG	GG	GG	CC	AA	AA	TT	AA	CC	AA	GG	TT	CC	GG	GG	AA	GG	GG	AA	AA	GG	CC	CC	TT	CC	TT	TT	TT
Upstr 4	AA	TT	AA	AA	AA	GG	CC	AA	GG	TT	AA	GG	CC	CC	AA	GG	CC	CC	AA	GG	GG	GG	AA	TT	GG	CC	AA	CC	CC
Upstr 5	AA	TT	GG	GG	CC	GG	CC	TT	GG	CC	AA	AA	CC	CC	GG	GG	AA	GG	AA	GG	GG	GG	CC	CC	GG	CC	TT	TT	TT
Upstr 6	GG	TT	AA	AA	CC	GG	CC	AA	GG	TT	AA	GG	CC	CC	AA	AA	AA	CC	AA	GG	GG	TT	AA	TT	GG	GG	AA	CC	CC
Upstr 7	AA	TT	AA	AA	CC	GG	CC	AA	GG	TT	TT	GG	CC	TT	GG	GG	AA	CC	AA	GG	GG	GG	AA	TT	GG	CC	TT	CC	TT
Upstr 8	AA	TT	..	..	CC	AA	CC	TT	GG	CC	AA	GG	CC	CC	GG	GG	AA	GG	AA	GG	GG	GG	CC	CC	GG	CC	TT	TT	TT
Upstr 9	AA	TT	AA	AA	CC	GG	CC	AA	GG	TT	AA	GG	CC	CC	AA	GG	AA	CC	AA	GG	GG	GG	AA	TT	GG	CC	AA	CC	CC
Upstr 10	AA	TT	AA	AA	AA	AA	CC	AA	GG	TT	AA	GG	CC	CC	AA	GG	CC	CC	AA	GG	GG	GG	AA	TT	GG	CC	AA	CC	CC
Upstr 11	AA	TT	AA	AA	CC	AA	CC	AA	GG	TT	TT	GG	CC	TT	GG	GG	AA	CC	AA	..	..	GG	AA	TT	GG	CC	TT	CC	TT
Upstr 12	AA	TT	GG	GG	CC	..	CC	TT	GG	TT	TT	GG	CC	CC	GG	GG	AA	CC	AA	GG	GG	GG	AA	TT	GG	CC	..	CC	CC

Haplotype Frequency	Indica	Aus	Trop Jap	Temp Jap	Aromatic	Admixed	Total	Average Shoot [Mo] per Haplotype	p-value	adj r^2
0.2247			1	16	2	1	20	4.59 [11]	-	-
0.1685		1		3	7	4	15	4.30 [6]	-	-
0.1236			11				11	4.81 [7]	-	-
0.1236	2	6			1	2	11	4.72 [5]	-	-
0.0787		7					7	5.16 [7]	**0.0579**	**0.0547**
0.0674			6				6	3.27 [3]	**0.0906**	**0.0397**
0.0674	5					1	6	2.80 [5]	**0.0039**	**0.1462**
0.0674	5	1					6	3.66 [3]	-	-
0.0449	4						4	3.95 [1]	-	-
0.0112		1					1	5.59 [1]	-	-
0.0112	1						1	4.13 [1]	-	-
0.0112	1						1	3.54 [1]	-	-

Figure 3.11 Sequence analysis of the CDS region and upstream regulatory region of *OsMOT1* . A) SNPs shown represent all SNPs in the gene model for which there is variation among resequenced and phenotyped lines. SNPs highlighted in light red are non-synonymous. SNPs highlighted in blue are found in the 3' UTR. SNPs inside the black rectangle are in the exon of *OsMOT1* , An '*' indicates that the SNPs are in the coding region of a putative flap endonuclease gene (LOC_08g01130) in the reverse direction on the complimentary strand, '+' indicates a SNP in the 3' UTR of LOC_08g01130, and SNPs highlighted in dark blue represent non-synonymous polymorphisms in the coding region of the overlapping gene. Haplotype numbers highlighted in blue are specific to the *Japonica* clade, those in light red are specific to the *Indica* clade. Genotypes in green are identical to the reference genome (Nipponbare), Yellow represents the alternate allele, and non-highlighted nucleotides represent missing data. The numbers in brackets preceding the average shoot molybdenum concentration indicate the number of resequenced and phenotyped lines with that haplotype that contribute to the phenotypic average. B) SNPs shown represent only SNPs upstream of *OsMOT1* that explained a significant percentage of the shoot Mo concentration in the resequenced lines. Asterisks indicate significance at the 0.05 (**) and 0.1 (*) levels. SNPs highlighted in blue are significant when only *tropical japonica* germplasm is considered. SNPs in orange are significant only when *indica* germplasm is considered. SNPs in yellow are significant when the entire population of phenotyped individuals is considered. SNP c8p85415 (unhighlighted) was not significant in this population, but is present in the HDRA and explains 25.3% of the shoot Mo variation in *aus* in the diversity panel.

Subpopulation-specific sequence analysis of the CDS region of *OsMOT1* revealed low levels of variation, particularly within the *Japonica* clade. The two *tropical japonica*-specific haplotypes ('CDS 1' and 'CDS 7') differed at only 3 SNPs in the 3'UTR and yet haplotype variation in *tropical japonica* explained 32% of the phenotypic variation for Mo shoot content (Figure 3.12 A). Considering that exonic variation across *OsMOT1* is low despite abundant phenotypic diversity, it can be concluded that it is unlikely that variation in the coding region of this candidate gene is driving natural variation for Mo shoot uptake.

Subpopulation-specific sequence analysis of the inter-genic upstream region proved to be much more useful in terms of explaining significant differences in shoot molybdenum content. Similar to the *tropical japonica* haplotypes for the coding region of *OsMOT1* , two upstream haplotypes were identified for this subpopulation (Figure 3.12 B). Twenty-three of the 28 SNPs upstream of *OsMOT1* explained a significant portion of the phenotypic variance for Mo shoot concentration among *tropical japonica* lines and were used to define upstream haplotypes. Among the resequenced and phenotyped lines, haplotype 'Upstream 1' was characterized entirely by reference alleles at all 23 loci in question and was found at very low frequency (~5%). The remaining 17 lines were split between haplotypes 'Upstream 3' and 'Upstream 6' in a nearly 2:1 fashion (frequencies of 66% and 33%, respectively). Haplotype 'Upstream 3' was characterized by reference alleles at 22/23 SNPs and explained 47% of the phenotypic variance (Average Mo: 4.8ppm). Similarly, haplotype 'Upstream 6' explained 35% of the phenotypic variance and was

characterized by non-reference alleles at 22/23 SNPs. Among the lines under consideration, 'Upstream 6' is always associated with the genic haplotype 'CDS 7' which is identified by the 3 non-reference alleles in the 3'UTR. The correlation between the strong phenotypic differences and the dramatically different regulatory haplotypes, combined with the high level of conservation in the exon of *OsMOT1* , supports the hypothesis that, at least in the *tropical japonica* subpopulation, expression levels of *OsMOT1* is likely the physiological driver of the observed phenotypic variation.

A

Haplotype Number	8:86771	8:86946	8:87741+	8:87784+#	8:87981*	8:88233*	8:88288*	8:88316*	Haplotype Frequency (in TRJ)	# TRJ lines	Average Shoot [Mo] per Haplotype (TRJ only) ppm	p-value	adj r^2
CDS 1	GG	CC	TT	TT	GG	GG	CC	AA	0.6667	12	4.67 [8]	**0.0512**	**0.3208**
CDS 7*	GG	CC	TT	CC	AA	CC	CC	AA	0.3333	6	3.27 [3]		

B

Haplotype Number	8:84506**	8:84536**#	8:84570**	8:84571**	8:84952**	8:84982**	8:85352**	8:85395**	8:85444**	8:85465**	8:85469**	8:85828**	8:85844**	8:85936**	8:85937**	8:85964**	8:85967**	8:86005**	8:86039**	8:86058**	8:86101**	8:86126**	8:86128**	Haplotype Frequency (in TRJ)	# TRJ lines	Average Shoot [Mo] per Haplotype (TRJ only) ppm	p-value	adj r^2
Upstream 1	AA	TT	GG	GG	AA	TT	AA	CC	TT	GG	GG	GG	GG	AA	AA	GG	CC	CC	TT	CC	TT	TT	TT	0.0556	1	3.64 [1]	-	-
Upstream 3	AA	GG	GG	GG	AA	TT	AA	CC	TT	GG	GG	GG	GG	AA	AA	GG	CC	CC	TT	CC	TT	TT	TT	0.6111	11	4.81 [7]	**0.0118**	**0.4707**
Upstream 6*	GG	TT	AA	AA	CC	AA	GG	TT	CC	AA	AA	CC	AA	GG	GG	TT	AA	TT	GG	GG	AA	CC	CC	0.3333	6	3.27 [3]	**0.0318**	**0.3524**

Figure 3.12 CDS and upstream haplotype structure of *OsMOT1* in *tropical japonica* among resequenced lines. A) Coding haplotypes of *OsMOT1* in *tropical japonica*. SNPs shown represent all SNPs in the gene model for which there is variation among phenotyped and resequenced lines. SNPs highlighted in light red are non-synonymous and in the exon of *OsMOT1* , SNPs highlighted in blue are SNPs found in the 3' UTR. SNPs inside the black rectangle are in the exon itself. An "*" indicates that the SNPs are in the coding region of a putative flap endonuclease gene (LOC_08g01130) in the reverse direction on the complimentary strand, "+" indicates a SNP in the 3' UTR of the overlapping gene, and SNPs highlighted in dark blue represent non-synonymous polymorphisms in the coding region of the overlapping gene. SNP c8p87784 (indicated with a # tag) is also present in the HDRA and explains 32% of the phenotypic variation for shoot Mo content among the *tropical japonica* lines in the diversity panel. B) Haplotypes present in the upstream region of *OsMOT1* that are present in *tropical japonica* germplasm. Only SNPs that explained a significant portion of the variation of shoot Mo concentration in *tropical japonica* were included (significance as indicated by the double asterisks). SNP c8p84536 (indicated with a # tag) is likewise present in the HDRA and explains 29.5% of the phenotypic variation among *tropical japonica* lines in the diversity panel.

Haplotypes 'CDS 2', 'CDS 3', and 'CDS 5' were all present in the *aus* subpopulation, and one ('CDS 5') was specific to *aus* and characterized by a non-synonymous SNP variant in the exon. However, none of the CDS haplotypes present in *aus* explained a significant portion of the phenotypic variance among *aus* lines (Figure 3.13 A). There were no SNPs in the upstream region that were significant for Mo shoot content, specifically among *aus* lines, so upstream haplotypes for this subpopulation were defined using only the 8 SNPs that were shown to be significant when regressed on Mo shoot phenotype data for the entire diversity panel. Among the 4 upstream haplotypes present in the *aus* subpopulation ('Upstream 4', '5', '8', and '10'), only two were found in appreciable frequency in *aus* ('Upstream 4' and '5'; Figure 3.13B). None of the regulatory haplotypes explained a significant amount of phenotypic variance among the resequenced *aus* lines, but one SNP (c8p84414 at MSU7 position 85415), while it did not show significance when regressed on the phenotype data for the re-sequenced lines, explained 25% of the phenotypic variation of Mo shoot content in the RDP1. The alternate allele for this SNP in *aus* is in perfect LD with haplotype 'Upstream 5' and 'CDS 5' which exhibit slightly (though not significantly) higher Mo concentrations in *aus* (5.2ppm vs. 4.4ppm). While this is markedly different from the circumstances in *tropical japonica*, it isn't surprising considering the fact that the GWA peak highlighting this candidate gene is very significant in the GWA analysis for ALL and for *tropical japonica*, but is not present in *aus*. The fact that a single SNP upstream of *OsMOT1* explains a significant difference in shoot Mo concentration of *aus* lines in the diversity panel, but not among the resequenced lines, could be explained by a low sample size (55 lines vs. 13 lines) among the resequenced lines.

Variation in the *indica* sub-population was similar to *aus* in that two out of the 4 CDS haplotypes present in the resequenced *indica* lines were defined by non-synonymous SNP variation in the exon that was specific to the *indica* subpopulation. Despite that, every haplotype present in *indica* likewise failed to explain a significant portion of the variance for Mo shoot uptake when only *indica* lines were considered (Figure 3.14A). Additionally none of the six haplotypes present in the *indica* subpopulation explain a significant amount of phenotypic variance (Figure 3.14B). It is notable however that one haplotype ('CDS 3'), which is differentiated from the reference genome at only two SNPs in the 3'UTR, has a frequency of 33% in resequenced *aus* lines (the high Mo accumulating subpopulation, average ~6ppm) and only 10% among the resequenced *indica* lines (the low Mo accumulating subpopulation, average ~3.5ppm). 'CDS 3' is also in perfect LD with 'Upstream 4' among *aus* lines and has the same frequency. Interestingly, the Indonesian *indica* accession "Sigadis" (NSFTV207) is the only *indica* that was both resequenced and phenotyped. It possesses these *aus*-like haplotypes and its shoot Mo concentration is almost double (6ppm vs. 3.5ppm) that of the other *indica* accessions. The fact that this one *indica* line possesses an *aus*-like promoter and 3'UTR and exhibits enhanced shoot Mo content relative to other indica lines provides evidence that even in *aus,* where *OsMOT1* may not be the primary driver of shoot Mo concentration, when introgressed into an *indica* background, exhibits significant phenotypic gains. Further, the significant phenotypic impact of the *aus* introgression in Sigadis provides evidence that it is the promoter region of the *OsMOT1* allele that impacts shoot concentrations of Mo.

A

Haplotype Number	8:86771	8:86946	8:87741+	8:87784+#	8:87981*	8:88233*	8:88288*	8:88316*	Haplotype Frequency (in *aus*)	# *aus* Lines	Average Shoot [Mo] per Haplotype (*aus* only) ppm	p-value	adj r^2
CDS 2	GG	CC	TT	CC	GG	CC	CC	AA	0.2000	3	4.65 [1]	-	-
CDS 3	GG	CC	TT	CC	GG	GG	GG	AA	0.3333	5	4.61 [4]	-	-
CDS 5*	GG	AA	TT	CC	GG	GG	CC	AA	0.4667	7	5.16 [7]	-	-

B

Haplotype Number	8:84570*	8:84571**	8:84771*	8:85395**	8:85415#	8:85828**	8:85967**	8:86005**	8:86126**	Haplotype Frequency (in *aus*)	# *aus* Lines	Average Shoot [Mo] per Haplotype (*aus* only) ppm	p-value	adj r^2
Upstream 5*	GG	GG	GG	CC	AA	GG	CC	CC	TT	0.4667	7	5.16 [7]	-	-
Upstream 4	AA	AA	GG	TT	GG	CC	AA	TT	CC	0.4000	6	4.37 [4]	-	-
Upstream 8	..	..	AA	CC	GG	GG	CC	CC	TT	0.0667	1	3.77 [1]	-	-
Upstream 10	AA	AA	AA	TT	GG	CC	AA	TT	CC	0.0667	1	5.59 [1]	-	-
p-value (*aus* Diversity Panel)					0.0051									
adj r^2 (*aus* Diversity Panel)					0.2530									

Figure 3.13 CDS and upstream, haplotype structure of *OsMOT1* in *aus* among resequenced lines. A) *aus* haplotypes in the coding region of *OsMOT1* among the resequenced population. The number in brackets following the average phenotypic value for each haplotype indicates the number of lines in the resequenced population for which phenotypic information is available. SNPs highlighted in light red are non-synonymous and in the exon of *OsMOT1* , SNPs highlighted in blue are SNPs found in the 3' UTR. SNPs inside the black rectangle are in the exon itself. An "*" indicates that the SNPs are in the coding region of a putative flap endonuclease gene (LOC_08g01130) in the reverse direction on the complimentary strand, "+" indicates a SNP in the 3' UTR of the overlapping gene, and SNPs highlighted in dark blue represent non-synonymous polymorphisms in the coding region of the overlapping gene. SNP c8p87784 (indicated with a # tag) is also present in the HDRA, but was monomorphic in *aus*. B) Haplotypes from the resequenced dataset in the upstream region of *OsMOT1* . Significance is similarly denoted with (*) or (**) and only SNPs that explained a significant portion of the variance when the entire population were considered. None of the SNPs were significant when only considering *aus* germplasm. SNP c8p85415 (indicated with a # tag) did not significantly explain any phenotypic variance in this population but did explain 25.3% of the phenotypic variance for shoot Mo content in *aus* in the diversity panel genotyped with the HDRA.

A

Haplotype Number	8:86771	8:86946	8:87741+	8:87784+#	8:87981*	8:88233*	8:88288*	8:88316*	Haplotype Frequency	# *indica* Lines	Average Shoot [Mo] per Haplotype (*indica* only) ppm	p-value	adj r^2
CDS 2	GG	CC	TT	CC	GG	CC	CC	AA	0.3000	6	3.74 [2]	-	-
CDS 3*	GG	CC	TT	CC	GG	GG	GG	AA	0.1000	2	6.10 [1]	-	-
CDS 4	TT	CC	TT	CC	GG	GG	CC	TT	0.3500	7	3.51 [4]	-	-
CDS 6	GG	CC	CC	CC	..	GG	CC	AA	0.2500	5	3.61 [2]	-	-

B

Haplotype Number	8:84570*	8:84571**	8:84660*	8:84771*	8:85395**	8:85399**	8:85450**	8:85465*	8:85778**	8:85828**	8:85936*	8:85967**	8:86005**	8:86101*	8:86126**	Haplotype Frequency	# *indica* Lines	Average Shoot [Mo] per Haplotype (*indica* only) ppm	p-value	adj r^2
Upstream 4*	AA	AA	AA	GG	TT	AA	CC	AA	CC	CC	GG	AA	TT	AA	CC	0.1111	2	6.10 [1]	-	-
Upstream 7	AA	AA	CC	GG	TT	TT	TT	GG	AA	CC	GG	AA	TT	TT	CC	0.2778	5	2.80 [4]	-	-
Upstream 8	..	..	CC	AA	CC	AA	CC	GG	AA	GG	GG	CC	CC	TT	TT	0.2778	5	3.61 [2]	-	-
Upstream 9	AA	AA	CC	GG	TT	AA	CC	AA	AA	CC	GG	AA	TT	AA	CC	0.2222	4	3.95 [1]	-	-
Upstream 11	AA	AA	CC	AA	TT	TT	TT	GG	AA	CC	..	AA	TT	TT	CC	0.0556	1	4.13 [1]	-	-
Upstream 12	GG	GG	CC	..	TT	TT	CC	GG	AA	CC	GG	AA	TT	..	CC	0.0556	1	3.54 [1]	-	-

Figure 3.14 CDS and upstream haplotype structure of *OsMOT1* in *indica* among resequenced lines. A) Coding haplotypes of *OsMOT1* in *indica*. SNPs shown represent all SNPs in the gene model for which there is variation among phenotyped lines. Highlighted in red are non-synonymous and in the exon of *OsMOT1* , SNPs highlighted in blue are SNPs found in the 3' UTR. SNPs inside the black rectangle are in the exon itself. An "*" indicates that the SNPs are in the coding region of a putative flap endonuclease gene (LOC_08g01130) in the reverse direction on the complimentary strand, "+" indicates a SNP in the 3' UTR of the overlapping gene, and SNPs highlighted in dark blue represent non-synonymous polymorphisms in the coding region of the overlapping gene. SNP c8p87784 (indicated with a # tag) is also present in the HDRA, but was monomorphic in *indica*. B) *indica* haplotypes from the resequenced dataset in the upstream region of *OsMOT1* . Significance is similarly denoted with (*) or (**) and only SNPs that explained a significant portion of the variance when the entire population (unhighlighted) or just the *indica* subpopulation (highlighted in orange) were considered.

Gene co-expression network analysis of *OsMOT1*

Further evidence supporting the role that gene expression of *OsMOT1* might play in determining shoot Mo concentration is found using gene co-expression network analysis. When *OsMOT1* is used as a single gene query in the riceFREND gene co-expression network database (Sato et al. 2013) a small network with 11 nodes and 10 edges is revealed (MR=7, Hierarchy=3). Intriguingly, two gene models (LOC_Os01g61070 and LOC_Os011g38790) not identified in the gene ontology-based candidate gene analysis were found to be co-expressed with *OsMOT1* (Figure 3.15) and are each found within 500kb of a GWA peak in our analysis for molybdenum root concentration in *temperate japonica* (Peak_IDs 94 and 95, respectively, in Supplementary Table 3.1). The specificity to *temperate japonica* is not surprising, considering the data that underlies riceFREND is collected largely on the *temperate japonica* reference accession 'Nipponbare.' It is also interesting regarding shoot Mo concentration to note that *temperate japonica* is the only sub-population for which there is no GWA peak in the region of *OsMOT1* for Mo shoot concentration, but there is a region of elevated significance on chromosome 11 that co-localizes with GWA peak #95, and with LOC_Os11g38790 (from the co-expression analysis). Given the low level of polymorphism in *OsMOT1* in the *temperate japonica* subpopulation, it is not unreasonable to expect that *OsMOT1* would not show significance in the GWA mapping analysis for that subpopulation, and that in a genomic background where the primary driver of shoot Mo relocation is fixed, other genes related more generally to whole-plant Mo accumulation might be significant. The fact that the two GWA peaks that co-localize with gene models co-expressed with *OsMOT1* are

significant only in a background where variation at *OsMOT1* is fixed might suggest that *temperate japonica* is a good target for investigating Mo accumulation mechanisms that may be underrepresented (or simply have a smaller phenotypic effect) in the other, more variable, rice gene pools.

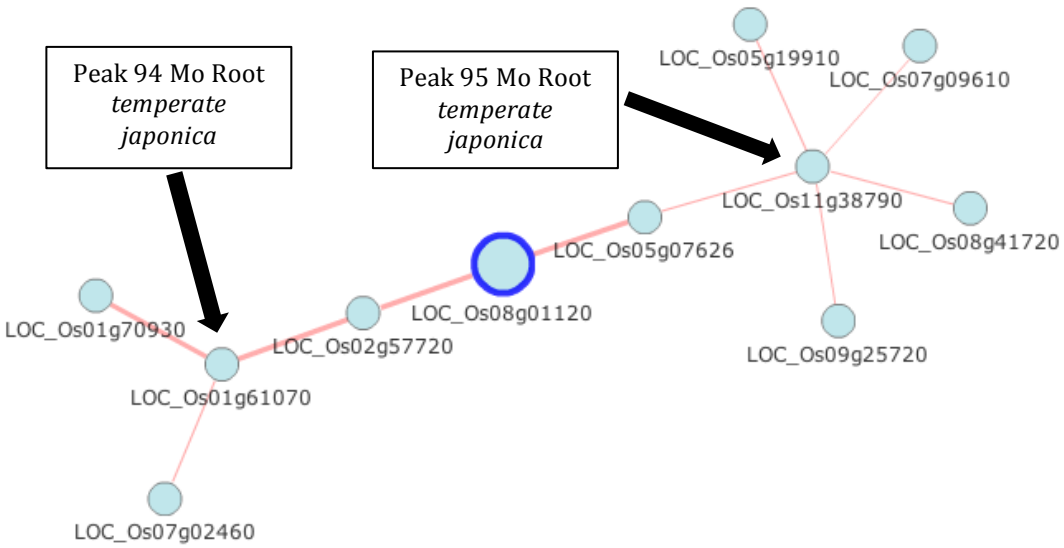


Figure 3.15 Gene co-expression network of *OsMOT1* . Results from a single gene search query using *OsMOT1* in the RiceFrend gene co-expression database (http://ricefrend.dna.affrc.go.jp/). The network (Hierarchy: 3, MR:7) has 11 nodes and 10 edges. LOC_Os01g61070 (annotated as an expressed heavy metal-associated domain containing protein) and LOC_Os11g38790 (annotated as an expressed protein) co-localize with peaks 94 and 95 in the GWA analysis for root Mo concentration in *temperate japonica*. Gene annotations from the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/index.shtml).

Molybdenum is required as a cofactor for nitrogen and sulfur metabolism in all plants, and plays an important role in abscisic acid biosynthesis. Despite it's essential role, very little is known about its uptake and regulation, nor about the genetics controlling these activities. Baxter et al. (2008) and Tomatsu et al. (2007) have elucidated some of the genetic control of molybdenum uptake and transport in *A. thaliana*, but nothing is currently known about similar mechanisms in rice or any other major crop plant. By mapping ionomic traits in rice, we implicate *OsMOT1* as one of the primary genes responsible for uptake of molybdenum into rice shoots and hypothesize that the mechanism driving the phenotypic variation is related to differential gene expression across lines.

Annotation of selected GWA peaks – Sodium

Next to the GWA peak for molybdenum shoot content, the next most significant peaks in the analysis are found in conjunction with sodium shoot concentration in *aus* (Peak ID 306, 311, and 314 respectively). Of the three, two are specific to the *aus* subpopulation (Peaks 306 and 311) and one peak (Peak ID 314) co-localizes with a similarly significant peak for sodium shoot concentration detected when the entire RDP1 is considered as a whole (Figure 3.16). The *aus* subpopulation is adapted to the Ganges river basin in northwestern India and Bangladesh and has contributed alleles for salt tolerance in the past (Thomson et al. 2010). The three peaks span 1.6, 0.6, and 0.59 megabases respectively and in each case the peak SNP has a negative LOG p-value of more than 8.4. There are 297 gene models in the Nipponbare reference genome under peak 306 on chromosome 1, none of which were identified using the

candidate gene search. However, one locus annotated as the small heat shock protein (sHSPs) *sodium and lithium tolerant 1* (*SLT1*) did co-localize with this peak and has been implicated in sodium transport in the past (Antoine et al. 2005, Matsumoto et al. 2001). Peak 311 located on chromosome 4 harbored only 85 genes underneath it. Among them include three widely studied high affinity potassium transporters (OsHKT1;1, OsHKT1;2, and OsHKT1;4; (Garciadeblas et al. 2003). In addition to the HKT transporters there was also one gene (LOC_Os04g51610) under the peak annotated as a sodium/calcium exchange p-type ATPase that was highlighted in the candidate gene search. Interesting, the three SNPs in this peak that were found to be within a gene model for a candidate gene were found in two of the introns (intron 6 and intron 33) of this candidate gene, and not in any of the HKT transporters in the region. The third peak (Peak 314) was located on chromosome 8 and included 113 genes. Among them, only two were identified in the candidate gene search (LOC_Os08g43690 and LOC_Os08g44310) with the first gene annotated as a member of the monovalent cation:proton anti-porter-2 family. The second ,candidate, LOC_Os08g44310, is not described in rice, but was highlighted by the candidate gene search because it is an ortholog of the *Z. mays* gene GRMZM2g067235, which is associated with the gene ontology term GO:0006816 describing mitochondrial sodium/calcium ion exchange. In addition to these three major peaks, there were nine other peaks present in the *aus* specific GWA analysis of sodium shoot content (Figure 3.17).

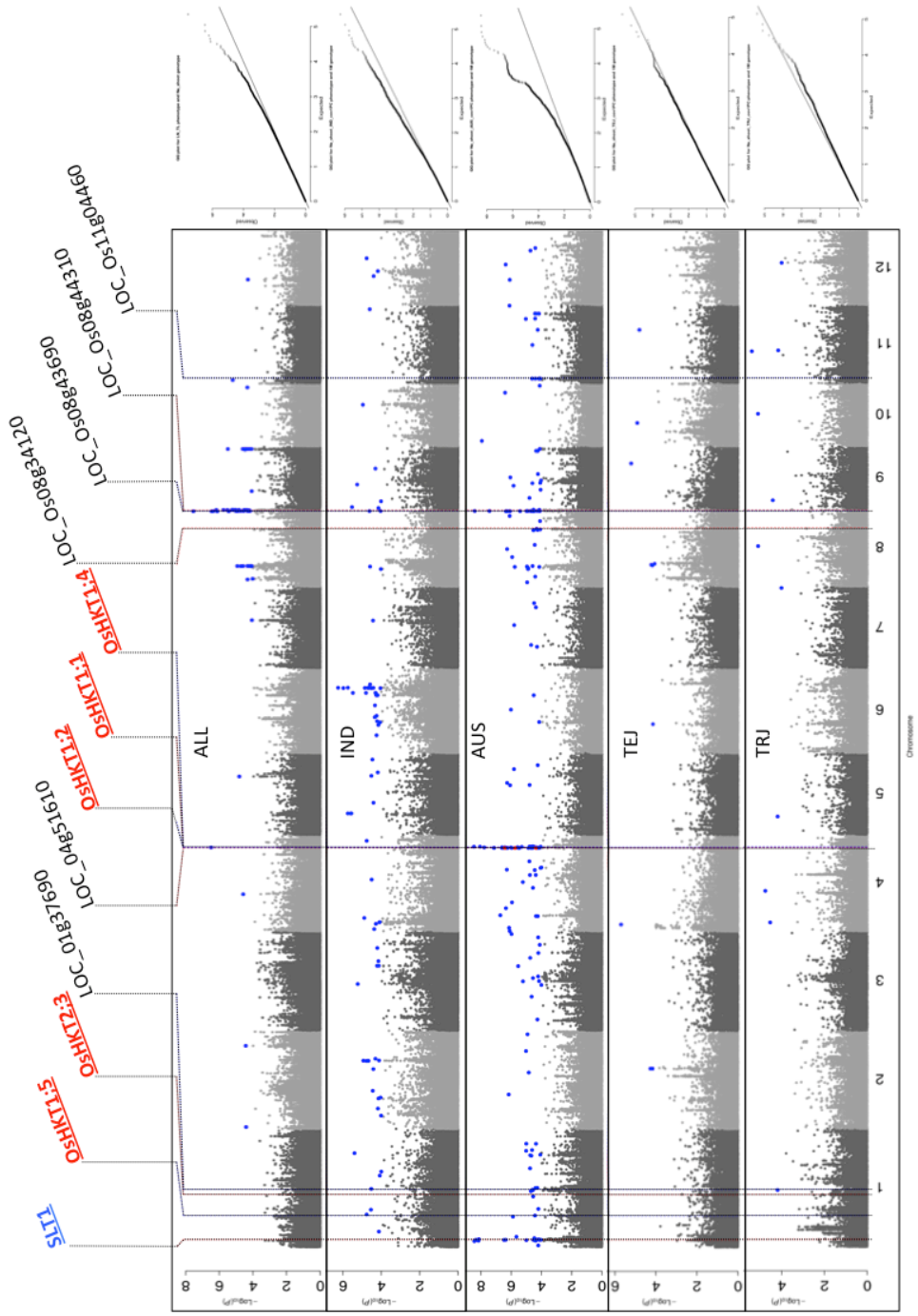


Figure 3.16 Manhattan and Quantile-Quantile plots for sodium shoot content. Blue SNPs have p-values less than 10^{-4}, and red SNPs are located in a candidate gene itself. Highlighted in red are the high-affinity potassium transporters co-localizing with peak 311 on chromosome 4 in the *aus* subpopulation, and in blue is sodium and lithium tolerant 1 co-localizing with peak 306 also in the *aus* gene pool.

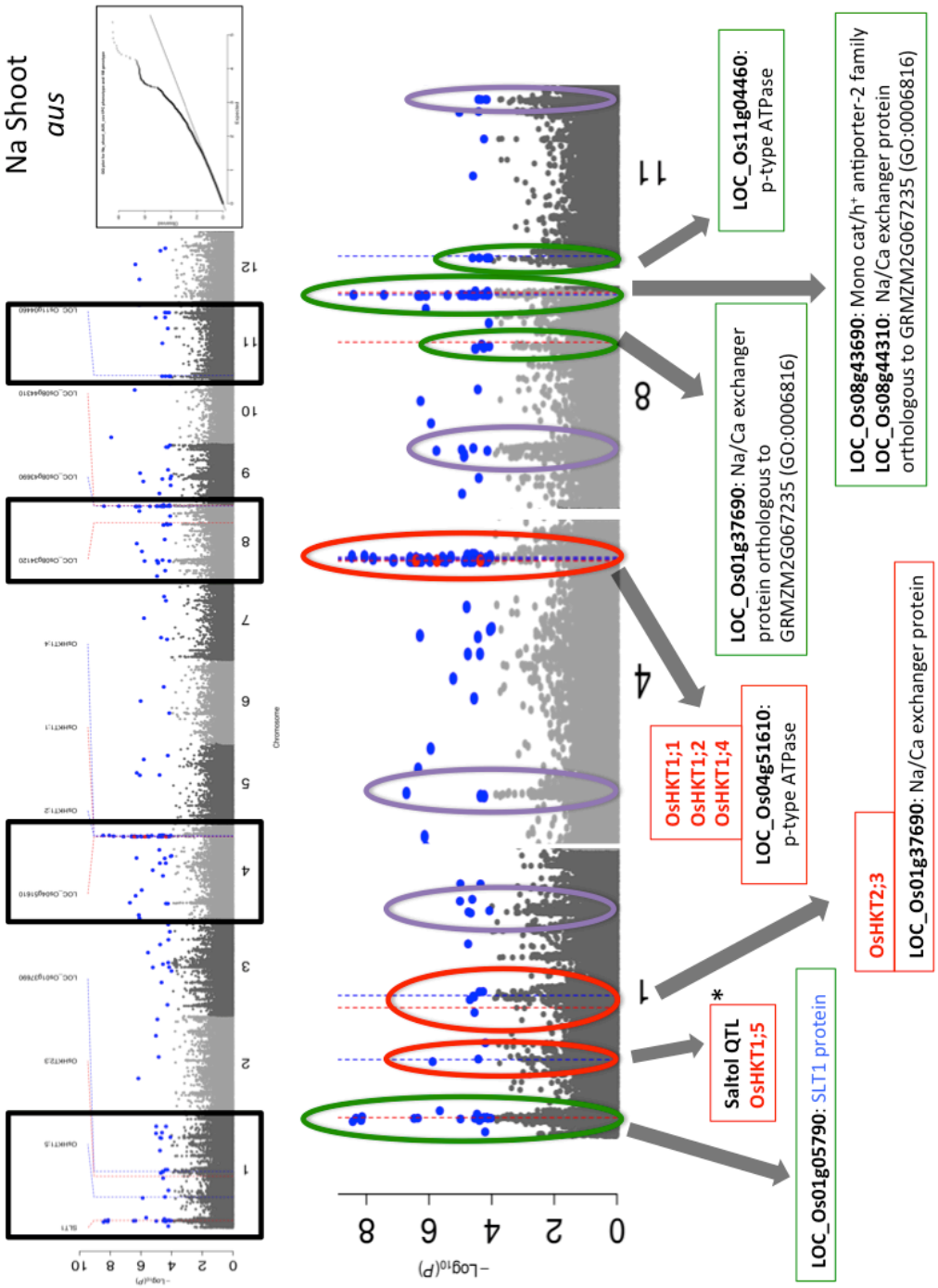


Figure 3.17 Overview of peaks for sodium shoot content in *aus*. Hits highlighted in red indicate peaks that co-localize with known rice sodium transporters, green with putative candidate genes, and purple where no candidates were detected. '*' indicates a the candidate genes are associated with a region of elevated significance that was not formally defined as a peak.

Haplotype analysis of *OsSLT1* using the HDRA

Sodium shoot concentration in the diversity panel was normally distributed, and exhibited statistically significant variation across subpopulations, but due to the significant variance still present within each subpopulation, the subpopulation identity only explained 3.5% of the phenotypic variation (p-value 0.0001; Figure 3.18). Notably, the mean of the *aus* subpopulation is lower than all the others, but even this gene pool has significant outliers. Noting that one of the most significant GWA peaks was found to co-localize with *OsSLT1* only in the *aus* analysis, and that *aus* has the lowest average shoot sodium content, is adapted to regions known for saline soils, and has donated alleles contributing to salt tolerance in the past, haplotype analysis was conducted at the *OsSLT1* gene to try to explain some of the Na shoot phenotypic variance (Figure 3.19). *OsSLT1* is a member of a recently discovered class of plant specific heat shock domain containing proteins similar to the small heat shock family of proteins (sHSPs). sHSPs-like proteins have been shown to be an important class of stress-related genes that have been demonstrated to be critical for response to temperature, water deficit, and osmotic balance (Antoine et al. 2005). The *SLT1* class of proteins was first discovered in *Nicotiana benthamiana* and *A. thaliana* via yeast mutant complementation tests, where Na sensitive mutant yeast strains harboring either the tobacco or *Arabidopsis* orthologs of *SLT1* showed enhanced tolerance to sodium and lithium toxicity via the CaN and SPK1/Hal4 signal transduction pathways. Interestingly, this gene has been shown to be constitutively expressed in rice, tobacco, and *Arabidopsis* (Sarkar et al. 2009) and Na/Li tolerance in yeast only occurred after

cleavage of the C-terminal domain of the *Arabidopsis* and tobacco proteins, implying a post-translational mechanism of auto-inhibition/activation (Matsumoto et al. 2001). The C-terminally truncated *OsSLT1* from rice has been shown to provide enhanced heat tolerance when heterologously expressed in recombinant *E. coli* and was shown to exhibit ATP independent molecular chaperone activity. Unlike the orthologs in *Arabidopsis* and tobacco, *OsSLT1* has not previously been implicated in sodium or lithium homeostasis (Antoine et al. 2005).

There were a total of nine SNPs from the HDRA in the *OsSLT1* gene model on the HDRA. Three of the SNPs are in the 5'UTR, three non-synonymous and two synonymous SNPs in the single exon of the gene, and one SNP is found in the 3'UTR. Using these 9 markers, nine distinct haplotypes were identified in the RDP1, the first five of which occurred with appreciable frequency in the population, and showed significant enrichment for one subpopulation or another (Figure 3.19). The *aus* subpopulation, in particular, had 4 distinct haplotypes, one of which ('Haplotype 3') displayed a markedly lower average sodium shoot content (260 ppm verses 297 ppm) and explained 3% of the phenotypic variance (p-value 0.0010) when the entire RDP1 was considered. There were other haplotypes (notably 'Haplotype 9') that were extremely rare in the population, but also exhibited comparatively low shoot sodium values. If that variation were indeed driven by genetic diversity at *OsSLT1*, it would not have been detected due to the low statistical power of GWA to detect the effects of rare alleles. Since only variation in *aus* showed a wide range of phenotypic classes, it is not surprising that only the analysis specific to the *aus* subpopulation would exhibit

a peak in this region. When only the *aus* lines are considered for the haplotype analysis, the two most abundant haplotypes in *aus* ('Haplotypes 3' and '5') show dramatically different levels of shoot sodium and explain 20% and 15% of the phenotypic variation, respectively (Figure 3.20). These two haplotypes are distinguished by variation for one SNP in the 5'UTR of *OsSLT1*, but are otherwise identical. There is variation at one of the non-synonymous SNPs in the *aus* gene pool, but the minor allele frequency is reasonably low and as such does not explain a significant amount of phenotypic variance.

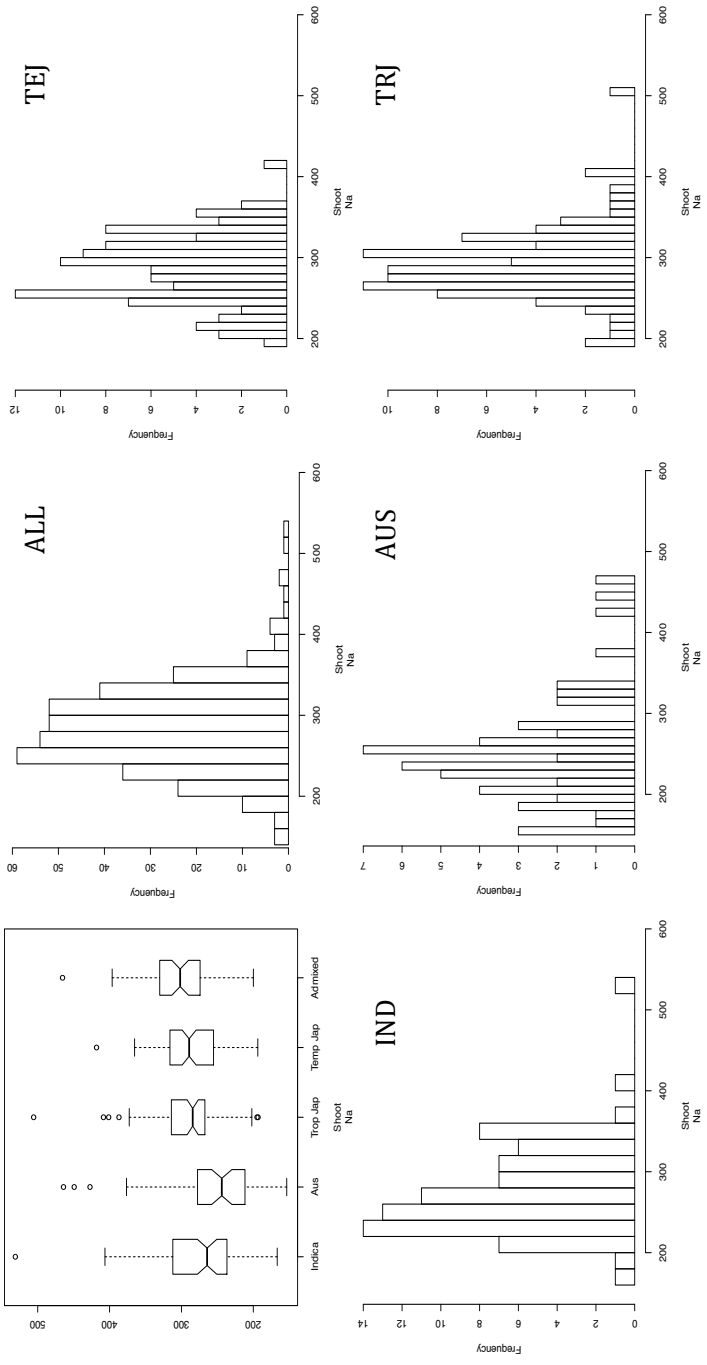


Figure 3.18 Boxplots and histogram distributions of sodium shoot phenotype data. Each subpopulation is represented individually, and ALL represents the entire panel. Subpopulation identity explains 3.6% of the phenotypic variation for Na shoot content (p-value = .000116).

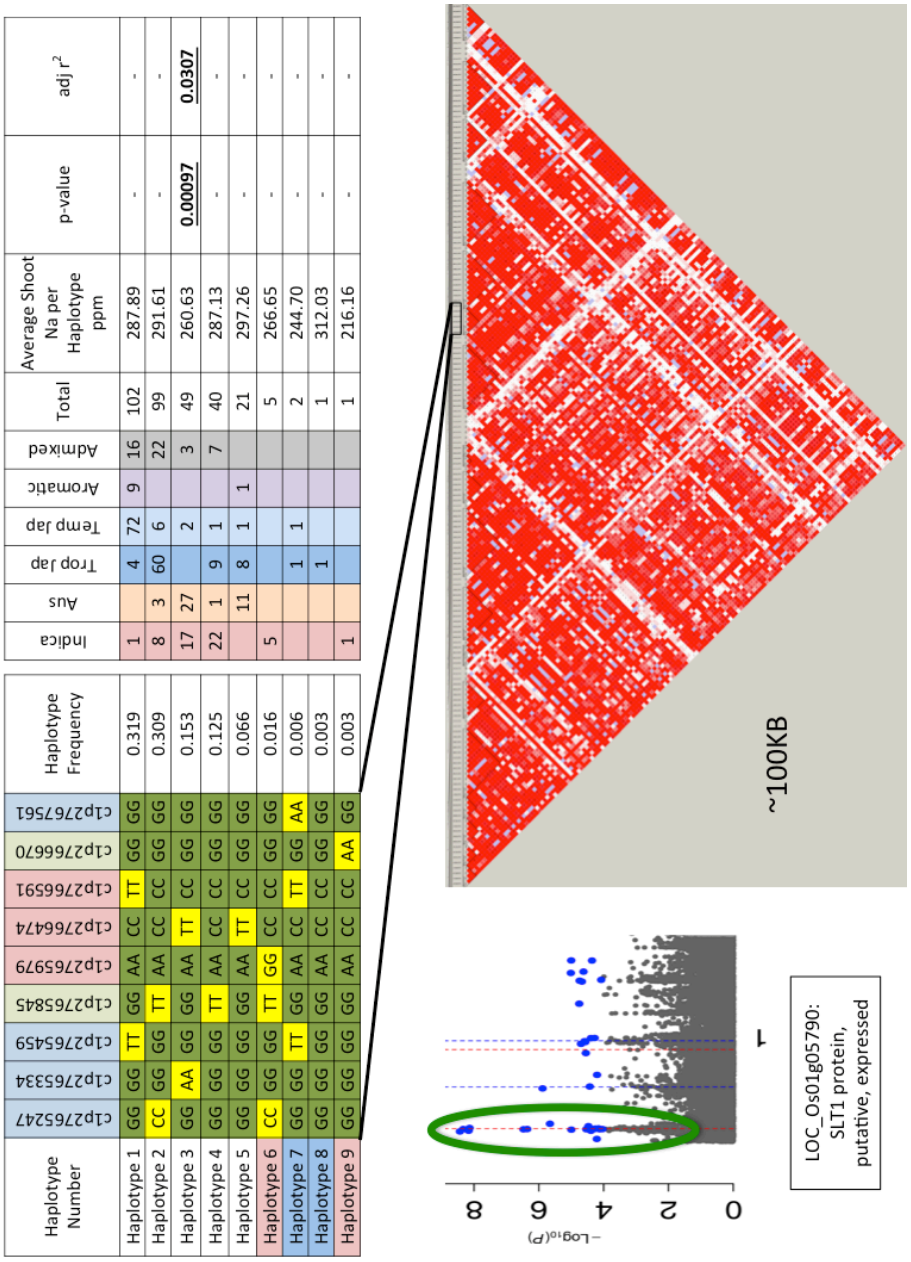


Figure 3.19 LD and Haplotype analysis of region surrounding *OsSLT1* . A) Haplotype and B) LD analysis of region surrounding *OsSLT1* based on genotype data from the HDRA on the diversity panel. SNPs in the haplotype analysis highlighted in light red SNPs are non-synonymous, light green are synonymous, and SNPs highlighted in blue are in the gene model but present in the 3'/5' UTR. Haplotype numbers highlighted in blue are specific to the *Japonica* clade, those in orange are specific to the *Indica* clade. Genotypes in dark green are identical to the reference genome (Nipponbare), Yellow represent the alternate allele. LD Plot spans the region from chr1:2715000 – chr1:2818000.

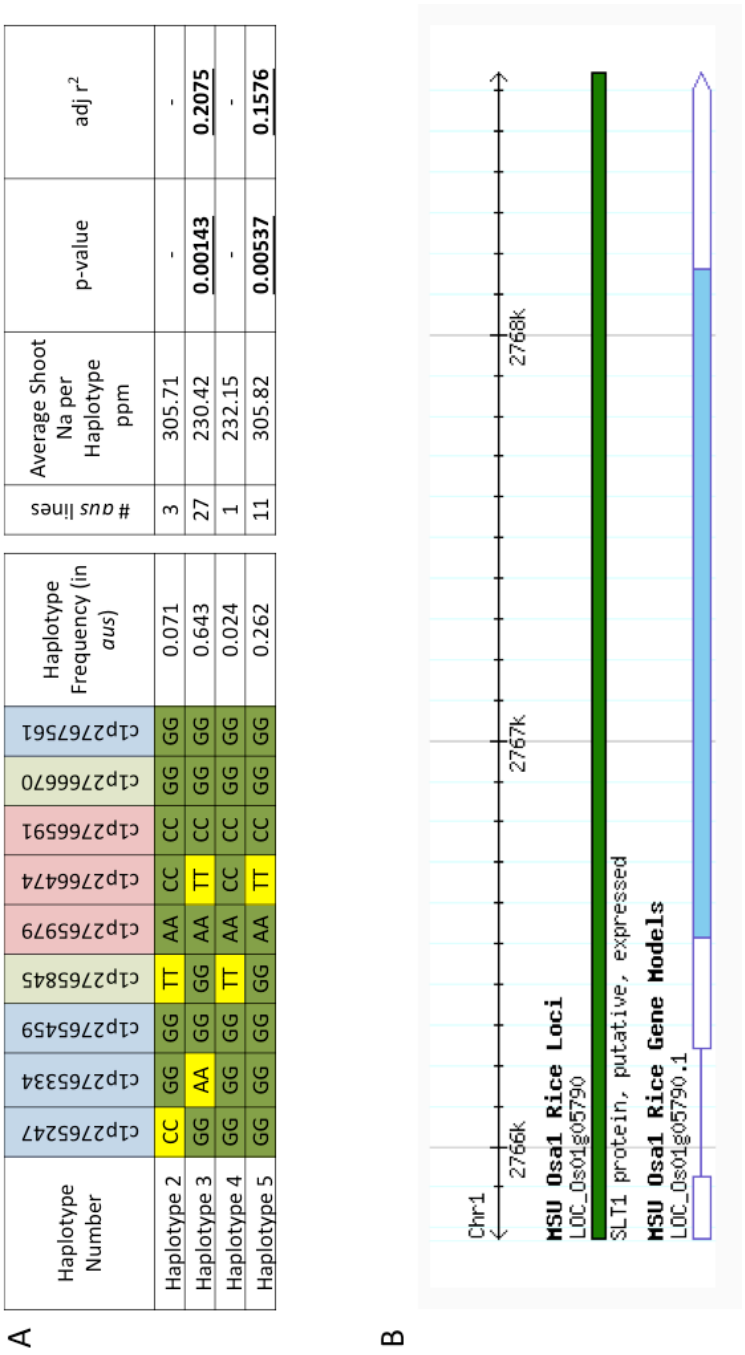


Figure 3.20 Subpopulation specific haplotype analysis of *OsSLT1* in *aus*. A) SNPs in the haplotype analysis highlighted light red SNPs are non-synonymous, green are synonymous SNPs highlighted in blue are in the gene model but present in the 3'/5' UTR. Genotypes in dark green are identical to the reference genome (Nipponbare), Yellow represent the alternate allele. B) Screenshot of the *OsSLT1* gene and transcript model from Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/index.shtml).

Sequence Analysis of *OsSLT1* using re-sequenced lines

Sequence analysis of *SLT1* was conducted using the set of resequenced lines described for the *OsMOT1* analysis in order to more fully explore the impact of genetic variation at this locus on Na shoot concentration. There were 14 total SNPs in the *OsSLT1* gene model: six in the 5'UTR, six in the exon (two synonymous and four non-synonymous) and two in the 3'UTR (Figure 3.21). Together these variants created seven distinct haplotypes across the entire population, with two specific haplotypes private to the *Japonica* clade ('CDS 2' and 'CDS 7'), one specific to the *Indica* clade, and one specific to the *aus* subpopulation ('CDS 6). Lines containing 'CDS 6', the *aus* haplotype, had markedly lower average shoot sodium content, and when regressed on the phenotype, explained 18% of the variation (p-value .001). When the analysis was repeated and limited to only the four haplotypes present in the *aus* lines ('CDS 1, 4, 5, and 6), a very different pattern emerged. Only haplotypes 'CDS 4' and 'CDS 6' occurred at frequencies >10% and while the phenotypic averages were quite different in the RDP1 as a whole, when only the *aus* lines were included in the comparison, the means of the two haplotype classes were not significantly different. Haplotypes 'CDS 5' and 'CDS 1' were rare among the *aus* lines (but more common among the higher accumulating *indica* and *tropical japonica* lines), but the *aus* lines with these haplotypes had phenotypic values markedly higher than the *aus* average (268 ppm and 335 ppm, respectively), more on par with or higher than the subpopulation averages of *indica* and *tropical japonica*. These two haplotypes are also distinguished from the more common *aus* haplotypes by variation at a non-synonymous SNP in the exon of *OsSLT1*.

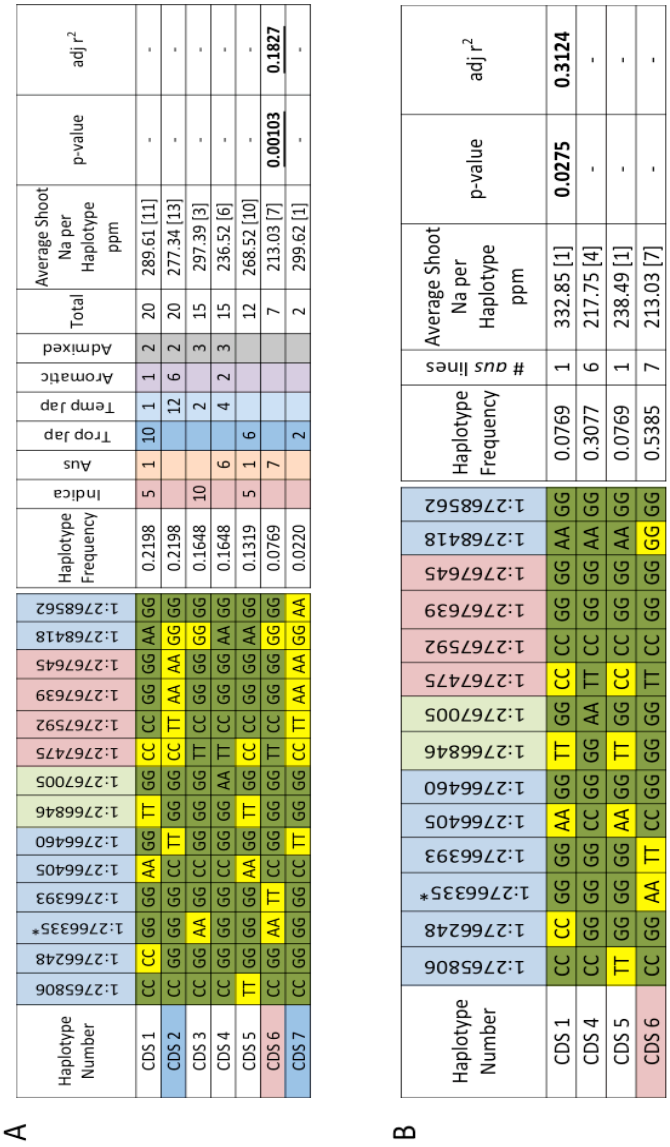


Figure 3.21 Sequence analysis of *OsSLT1* . A) SNPs shown represent all SNPs in the gene model for which there is variation among resequenced and phenotyped lines. SNPs highlighted in light red are non-synonymous, green are synonymous, and SNPs highlighted in blue are found in the 5'/3' UTR. An '*' indicates the SNP that was likewise present on the HDRA and distinguished the two *aus* haplotypes that explained 21% and 16% of the phenotypic variance each. Haplotype numbers highlighted in blue are specific to the *Japonica* clade, those in light red are specific to the *Indica* clade. Genotypes in green are identical to the reference genome (Nipponbare), Yellow represents the alternate allele. The numbers in brackets proceeding the average shoot sodium concentration indicate the number of resequenced and phenotyped lines with that haplotype that contribute to the phenotypic average. B) Sequence analysis of *OsSLT1* limited only to the *aus* lines in the population.

Analysis of *OsSLT1* with near-isogenic admixed lines

It is clear that patterns of haplotype variation at *OsSLT1* are consistent with the GWA analysis and this region of the genome is significant for shoot Na accumulation in the diversity panel when grown in hydroponics . In order to investigate the effect that the *aus* alleles have on salt tolerance under stress, near-isogenic lines (NILs) developed from a cross between the *tropical japonica* variety 'Jefferson' and the *aus*-like wild rice *Oryza rufipogon* (accession IRGC105491) were screened under salt stress. Five distinct lines were available for analysis: the two parental genotypes and three NILs, each backgrossed 3 times to the *tropical japonica* parent, and sharing the same background introgressions from the wild donor variety. The *tropical japonica* parent carries haplotype 'CDS 5' that is rare in *aus* but common in *indica* and *tropical japonica*, and the *O. rufipogon* donor line carries haplotype 'CDS 6'. This haplotype is specific to the *aus* subpopulation, and distinguished from 'CDS 5' by a non-synonymous SNP in the exon, a synonymous SNP in the exon, and 5 SNPs in the 5'/3' UTR of the *OsSLT1* gene.

The three NILs differed from each other primarily in the region surrounding *OsSLT1* , and had very similar background introgressions in the rest of the genome (Figure 3.22; Supplementary Table 3.4). Line 141_2-1 harbored *O. rufipogon* alleles from 0.6 to 20.2 megabases on the distal end of the short arm of chromosome 1. This introgression covers the *OsSLT1* region as well as a region containing the OsHKT1;5, OsHKT2;3, and a previously identified QTL for salt tolerance known as the Saltol QTL (Thomson et al. 2010). Line 141_2-6 harbors *O. rufipogon* alleles at the SalTol/HKT region, but

harbors *tropical japonica* alleles in the region of *OsSLT1* , specifically the 'CDS 5' haplotype of *OsSLT1* . Thus comparisons between these two lines should highlight the effect of the *OsSLT1* region on the phenotype. Line 141_2-11 harbors only 'Jefferson' alleles on chromosome 1, but has very similar introgressions of *O. rufipogon* in the rest of the genome and was included to estimate the effects. Five replicates each of treatment and control plants were grown under controlled conditions and bottom watered with either elevated sodium (EC of 18 using NaCl) or standard modified Magnavaca's solution (Famoso et al. 2010) at pH 5.8.

Figure 3.22 illustrates both the Na shoot concentration of each line and the Na/K ratio in the shoot tissue (a well-known proxy for salt tolerance in the field). Low Na/K ratios are an indication of less displacement of potassium ions by sodium ions under salt stress and indicate a tolerant genotype (Asch et al. 2000). The parental genotypes 'Jefferson' and *O. rufipogon* had differing but statistically indistinguishable shoot sodium concentrations (Dunnett Multiple Comparisons (DMC) test adjusted p-value: 0.6224) and shoot Na/K ratios (DMC adjusted p-value: 0.3268). However, the shoot sodium concentration and Na/K ratio for line 141_2-1, which differs from the other two introgression lines because of the introgression of the *aus*-like allele for *OsSLT1* on chromosome 1, were significantly lower than the shoot Na concentration and Na/K ratio for line 141_2-6 (DMC adjusted p-value: 0.0081). These differences are evidence that, at least in the genetic backgrounds of these NILs, the *aus*-like allele in the *OsSLT1* region alters shoot sodium accumulation and is likely to contribute to salt tolerance in an otherwise susceptible background. Likewise, 141_2-11 which acted as

a control for the effects of the remaining *O. rufipogon* introgressions in the genetic background was indistinguishable from the 'Jefferson' recurrent parent in shoot Na content and Na/K ratio (DMC adjusted p-value: 0.8229 and 0.9615 respectively), indicating that the background introgressions had little effect on the Na/K ratios or on Na content in the shoot. This demonstrates that the *aus*-like alleles in the *OsSLT1* region contributed by the wild donor significantly alter the Na/K shoot ratios of the NILs. As such, we hypothesize that variation at this locus can impact performance under salt stress and that the *aus* subpopulation is likely the best source of salt tolerant haplotypes.

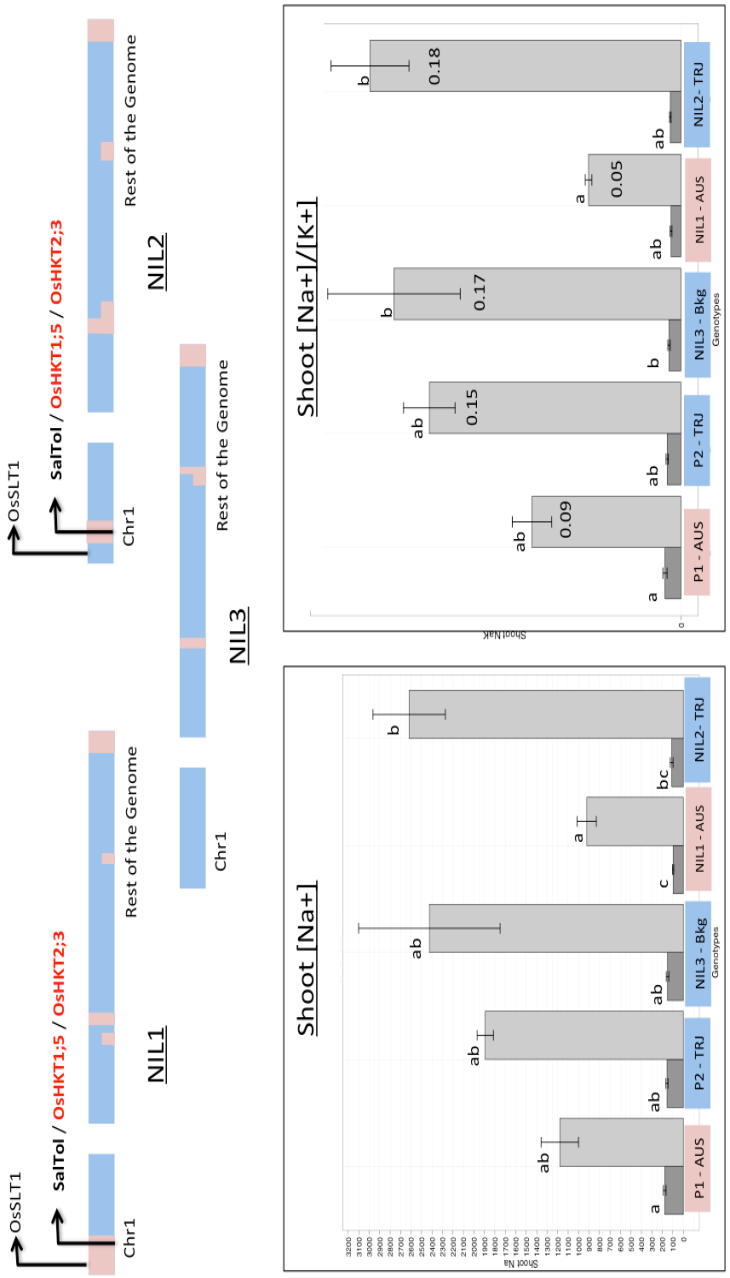


Figure 3.22 Summary of Na shoot content and shoot Na/K ratio for the bi-parental admixed lines. A) Graphical schematic of *O. rufipogon* introgressions of admixed lines and their relative position to genes/QTL of interest including *OsSLT1* . B) Bar plots of Na shoot content (left) and shoot Na/K ratio (right). Darker bars represent plants grown under control (non-sodium stress) conditions, and lighter bars represent plants grown under elevated sodium levels. Statistical groupings were determined using a Dunnett Multiple Comparison test for groups with unequal variance. Numbers in the bars of the right graph indicate the average Na/K ratio for each line. Error bars indicate the standard error of the mean.

CONCLUSION

This study represents the most comprehensive survey of ionomic variation to date. It is the first to provide evidence that *OsMOT1* controls the acquisition and partitioning of molybdenum to rice shoots and to propose a mechanism for this partitioning based on patterns of natural genetic variation. This is likewise the first report to highlight the role of natural genetic variation at the *SLT1* locus as a determinant of shoot sodium concentration and a possible selection target for increased salt tolerance in rice. We demonstrated that combining the power of controlled environments and Bayesian modeling can produce precise phenotype data for GWA analysis, which when combined with the otherwise disparate lines of evidence, like QTL mapping, candidate gene searching, and haplotype analysis, is an efficient way to develop targeted and coherent hypotheses about genes that underlie complex trait variation. Many of these gene targets are worthy of continued investment in the form of molecular physiological evaluation and integration into breeding populations. By navigating from phenotypic data collection through GWA analysis and a focus on haplotype variation to identify putative candidate genes, we illustrate the power of combining large genotypic and phenotypic datasets from shared germplasm resources to generate hypotheses. This work provides insight into the potential for rapid genetic gain that is possible when alleles from different subpopulations are introgressed into novel genetic backgrounds to make significant gains in phenotypic performance. Additionally, the use of NILs to investigate the potential of candidate genes represents one way in which specific and genetically useful germplasm can be utilized in a targeted way to test hypotheses generated by large genomic screens of natural variation.

MATERIALS AND METHODS

Plant Growth Conditions and Germplasm – Ionomics Analysis

Rice seeds were sterilized with 20% bleach for 15 to 20 minutes and germinated in rolled germination paper at 26 to 30 degrees C for 3-5 days under dark conditions. Upon germinating, seedlings were transferred to one of five 400 liter hydroponics tubs by suspending them in foam circles embedded in modified 35mm film canisters and placed in a foam framework that accommodated 200 plants per tub. Individuals were randomized within and across tubs. Aeration was not provided. Plants were stagger germinated such that five tubs were planted over the course of 5 days and likewise tissue was harvested over 5 days. The seedlings were allowed to grow in the hydroponic nutrient solution for 42 days before harvesting root and shoot tissue for ICP analysis. After the first 21 days, ribbon scaffolding was erected to support the shoots and prevent lodging. The nutrient solution was specially constructed and consisted of 1000 µM $Ca(NO_3)_2.4H_2O$; 1000 µM NH_4NO_3; 250 µM KNO_3; 250 µM KH_2PO_4; 250 µM $MgSO_4.7H_2O$; 1000 µM MES pH buffer; 77 µM $Fe(NO_3)_3.9H_2O$; 77 µM H_3EDTA; 300 µM NaOH; 50 µM KCl; 12.5 µM H_3BO_3; 0.1 µM $Na_2MoO_4.2H_2O$; 2 µM Na_2SeO_3; 4 µM $Na_2HAsO_4.7H_2O$; 0.1 µM KI; 0.1 µM $K_2Cr_2O_7$; 0.1 µM $NiSO_4.6H_2O$; 2 µM $MnSO_4.H_2O$; 1 µM $CuSO_4.5H_2O$; 0.1 µM $Co(NO_3)_2.6H_2O$; 2 µM $ZnSO_4.7H_2O$; 0.1 µM $Cd(NO_3)_2.4H_2O$; 0.1 µM $RbNO_3$; 0.1 µM $Sr(NO_3)_2$; 0.1 µM $LiNO_3$; 0.1 µM $PbNO_3$; and 145µL/L of commercial AgSil 25 (http://www.pqcorp.com/) which contributed 600 µM K_2SiO_3. Chemical speciation was determined according to Shaff et al. (2010) using the GEOCHEM-EZ speciation program and the predicted ionic activity (availability) for each element was

determined through the primary distribution and case progress table output. The solution in each tub was re-circulated daily with a magnetic drive pump and brought up to a pH of 6.0 using KOH. 15mL samples of the nutrient solution were taken every Monday, Wednesday, and Friday and used to evaluate the nutrient composition of the media. Given the volume of the tubs and the expense of the DI water used to fill them rather than changing the solution after a specified period of time, nutrients that depleted quickly were replenished as needed. Plant roots were desorbed in 5 mM $CaCl_2$ for 15 minutes and then the entire plant was rinsed in DI water prior to separating the roots from the shoots. Plant tissue was dried in a hot air drying oven at 60 degrees C for 5 days. The growing conditions were set to 30 degree C daytime temperatures and 26 degree C night time temperatures on a 12H light/dark cycle. Light intensity was approximately 450 mmol photons m^{-2} s^{-1}. The 394 genotypes used in the association analysis are part of a set of 410 *O. sativa* lines that have been genotyped with 44,000 SNPs (Tung et al. 2010) and again with the HDRA (McCouch S, personal communication). The experiment was repeated twice (with two reps of each line each time) in order to compile the 4 reps that comprised the raw data that went into the Bayesian adjustment pipeline.

Plant Growth Conditions and Germplasm – Salt tolerance screen

Plant growth chamber conditions were as above. Seeds from each line were sterilized as above and germinated in 6" pots containing the potting soil "Cornell Mix." Seven pots were then set in a tank and bottom watered with either modified Magnavaca's solution(Famoso et al. 2010) at pH 5.8 or the same solution amended with NaCl.

Treatment plants were allowed to grow to for 3 days with no salt stress, then for 4 days at 12 dS/M of NaCl stress, and finally for 14 additional days (for a total of 21 days) at 18 dS/M of NaCl. The nutrient solution in the tanks was replaced once a week. Shoot tissue was harvest and analyzed as above via ICP analysis to determine sodium content.

ICP Analysis for elemental composition

Dried shoot tissue was weighed and homogenized in liquid nitrogen and 200 mg were subsampled for ICP analysis. As whole roots were analyzed, they were not homogenized prior to weighing and digestion. Each tissue sample was digested in a carbon heat block using a Vulcan 84 automated sample digestion unit (http://www.qtechcorp.com/). 250 µL of 40ppm Yttrium standard was added to each sample in order to back-calculate the appropriate dilution volume. Approximately 4 ml of 60/40 nitric/perchloric acid was added to each sample and heated to 150 degrees C. After incubating for an hour at 150^{o} C, the blocks were heated to 190^{o} C for 10 minutes. Samples were not allowed to digest to dryness and were subsequently diluted with deionized 18 Mohm water. ICP analysis was run on a Thermo 6000 series inductively coupled argon plasma optical emission spectrometer housed at the US Department of Agriculture – ARS Robert W. Holley Center in Ithaca, NY.

Bayesian adjustment of phenotype data

Bayesian models were implemented using the open source statistical package R (http://www.r-project.org/). Replicates were modeled with a multivariate Student-*t*

distribution with three degrees of freedom to modulate the influence of outlier observations. Multivariate Gaussian distributions were assumed for location parameters at the rest of the levels. Three technical replicates of each ion measurement were collected by the ICP for each sample. The %RSD was treated as a calculated measurement error from these technical replicates and used to weigh individual observations within the hierarchical model, as described by (Greenberg et al. 2011). We noted four sets of replicate-level conditions with the potential to affect systematically our observations: the position of each plant within the growth chamber, the tub identity, the identity of the individual harvesting the tissue, and the identity of the digestion batch as each sample was prepared for ICP analysis. We constructed dummy variables reflecting contrasts within each set of covariates, but some of the resulting contrasts were collinear. To circumvent this problem, we estimated principal components of the corresponding covariance matrix and used principal component vectors that corresponded to non-zero eigenvalues as predictors in a variable-intercept regression at the replicate level in our model. To obtain biologically interpretable data-points, the levels of each element in each sample were divided by the weight of that sample. While this procedure of dividing by sample weight is preferable in highly structured experiments such as ours, because whole roots (rather than subsamples) were analyzed due to their smaller size, results have to be interpreted with the possibility of spurious correlations among variables in mind, particularly for the root samples. Only a portion of each shoot was used, resulting in a very narrow distribution of sample weights. A detailed discussion of the pitfalls and advantages of per-weight measurements can be found in (Greenberg et al. 2011).

Mixed-Model Genome-wide Association Analysis

Genome-wide association analysis was performed using the EMMAX mixed model approach to control for population structure. The model can be written as $y = X\beta+C\gamma+Z\mu+e$ where β and γ are coefficient vectors for SNP effects and subpopulation PCs respectively and treated as fixed effects. X and C are the SNP vector and the first 4 PC vectors and e is random error. The coefficient μ in this instance is the random effect accounting for differential relatedness between lines and Z is the corresponding design matrices. We assume $\mu \sim N(0,\sigma^2_g K)$ and $e \sim N(0,\sigma^2_e I)$, and K is the IBS matrix as in (Atwell et al. 2010). For within subpopulation analyses the model was changed to a simple regression $y = X\beta+Zu+e$ since no main population structure needed to be accounted for.

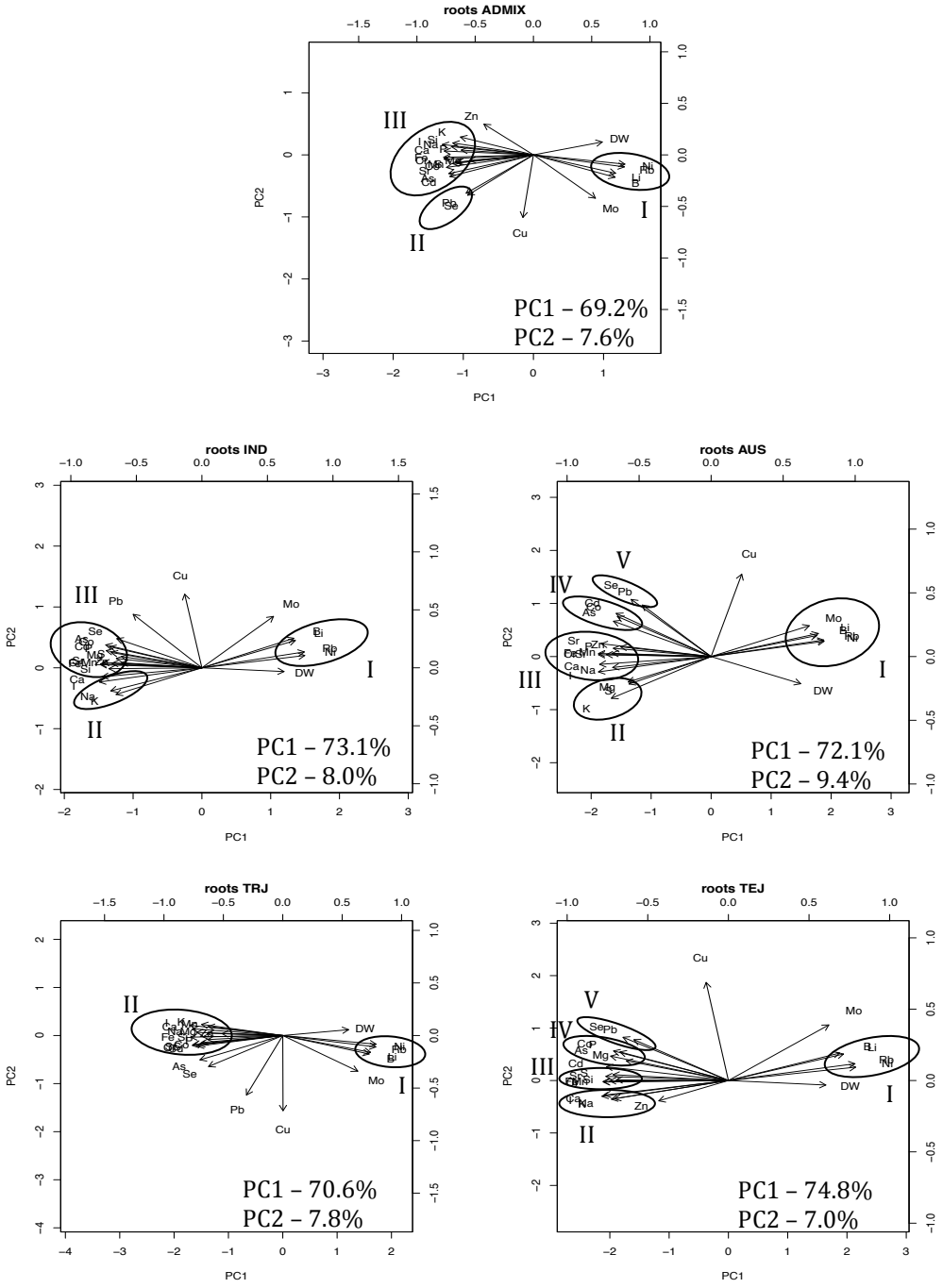
LD and Haplotype Analysis

Linkage disequilibrium analysis and haploblocks were calculated at the regions specified using the publically available Haploview software package(Barrett et al. 2005).

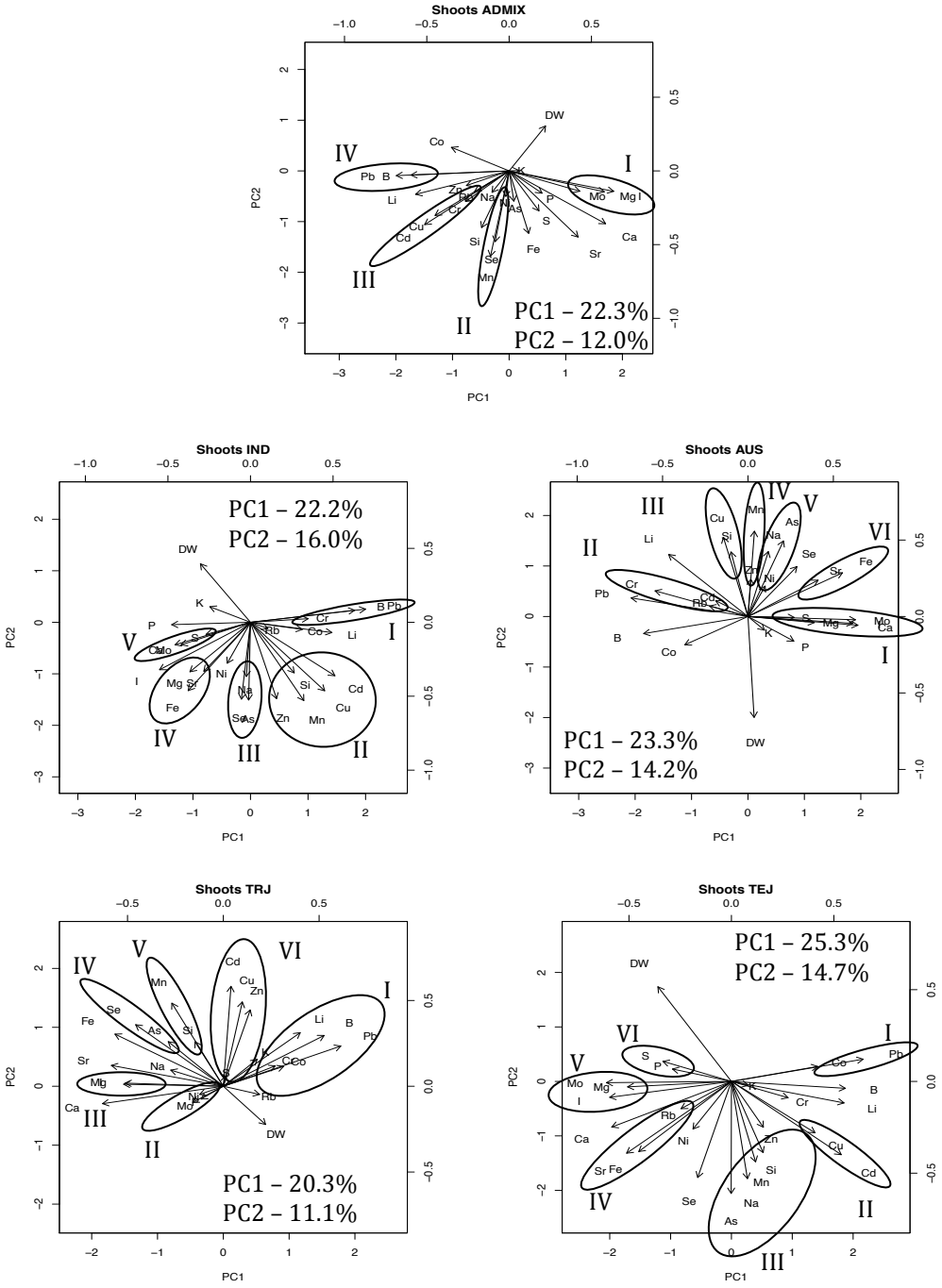
APPENDIX

Subpopulations specific analysis of PCA Biplots of the root and shoot ionome.

The relative interconnectedness of the ionome becomes more interesting from a breeding and genetics standpoint when we consider how these relationships change across subpopulations (Supplementary Figures 3.1 and 3.2). For the root ionome, the observed relationships in each subpopulation are not substantially different from the plot considering the entire diversity panel. With the exception of an arbitrary inversion of the axis along the second PC the relationships between each element and between the ionome and root DW do not change when considered across subpopulations.



Supplementary Figure 3.1 PCA Biplots of ionomic relationships within the roots by subpopulation.



Supplementary Figure 3.2 PCA Biplots of ionomic relationships within the shoots by subpopulation

Similar to the subpopulation specific relationships among the root ionome, phenotypes in the shoot ionome also do not fluctuate wildly across subpopulations. There are however, notable differences in the strength of the correlation between certain elements that does change with subpopulation. For example, among the *aus* lines there are much fewer instances of positive correlation with whole shoot biomass. This could be important to consider in light of the dwarfing genes from the green revolution that dramatically altered the harvest index of cultivated rice. Consistent with relationships depicted in Supplementary Figure 3.2 reducing shoot biomass in *indica* or *temperate japonica* might lead to marked increases in shoot concentrations of micronutrients/heavy metals, including Zn, Cu, Mn, and Cd (see clusters IND(II); TRJ(V); and TRJ(VI), where such metals occur in the soil. In the *aus* and *tropical japonica* subpopulations, semi-dwarf varieties would likely also accumulate high concentrations of the same heavy metals, with the exception of Cd (see clusters TEJ(II); TEJ(III); AUS(II); AUS(III); and AUS(IV)) because the correlative relationship between Cd and shoot dry weight in those subpopulations is dramatically reduced compared to *indica* and *temperate japonica*.

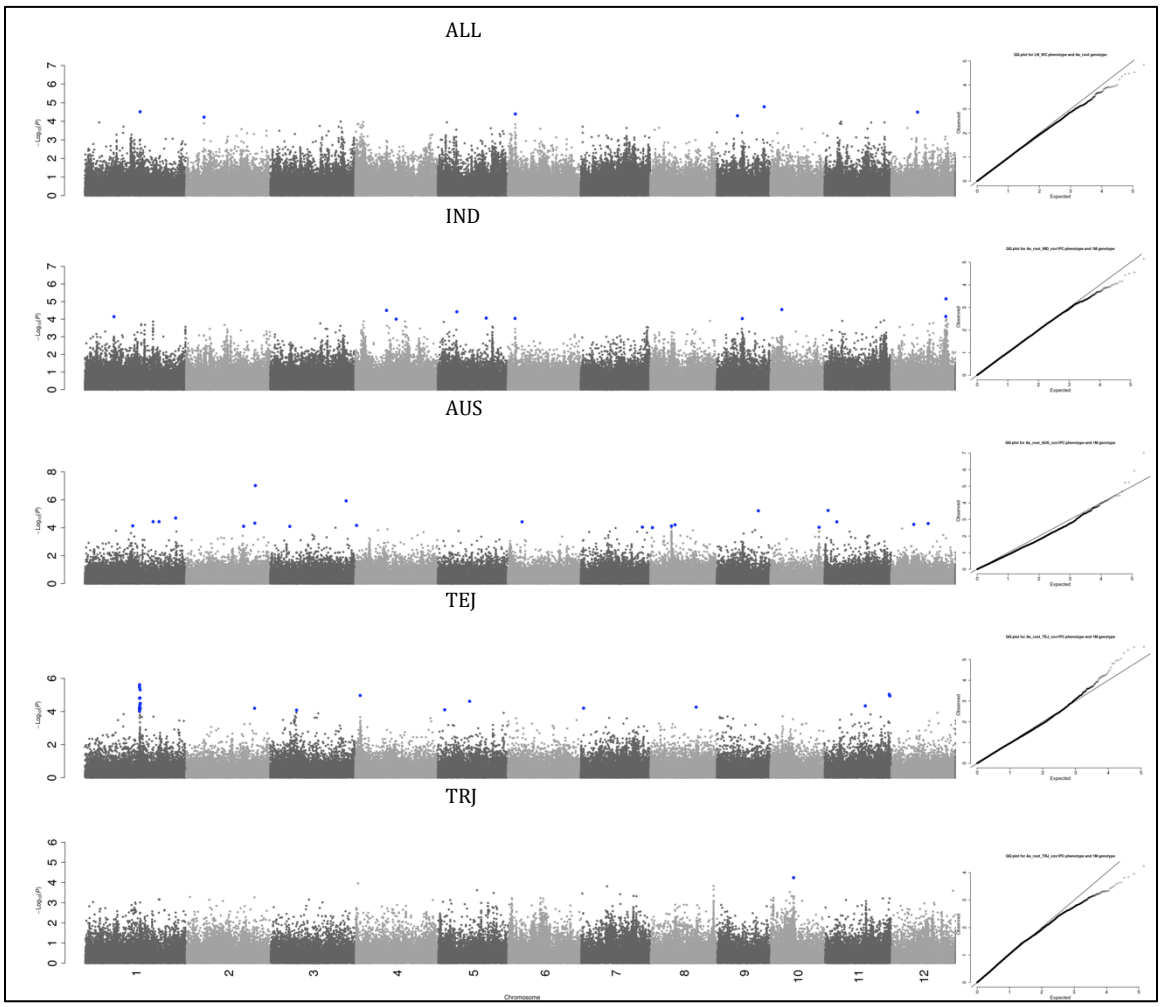
Similarly, Supplementary Figure 3.2 suggests that remobilization of both Cd and Cu to the shoots is tightly linked in all subpopulations except *aus* (see clusters AUS(II) and AUS(III)), and that P, S, As, and Zn concentrations, are positively correlated in all subpopulations except *tropical japonica* (where As and Zn share almost no relationship and P and S are slightly negatively correlated; see clusters TRJ(II); TRJ(IV); and TRJ(VI)). For some elemental pairs (notably Na/As [clusters AUS(V),

IND(III)], Cu/Si [clusters AUS(III), IND(II)] and Zn/Mn [clusters AUS(IV), IND(II))], the *Indica* clade harbors more strongly correlated relationships, while the same correlations in the *Japonica* clade are still present, but much less predictive of each other (Na/As [clusters TRJ(IV), TEJ(III)], Cu/Si [clusters TRJ(V), TRJ(VI), TEJ(II), TEJ(III)], Zn/Mn [clusters TRJ(V), TRJ(VI), TEJ(III)]).

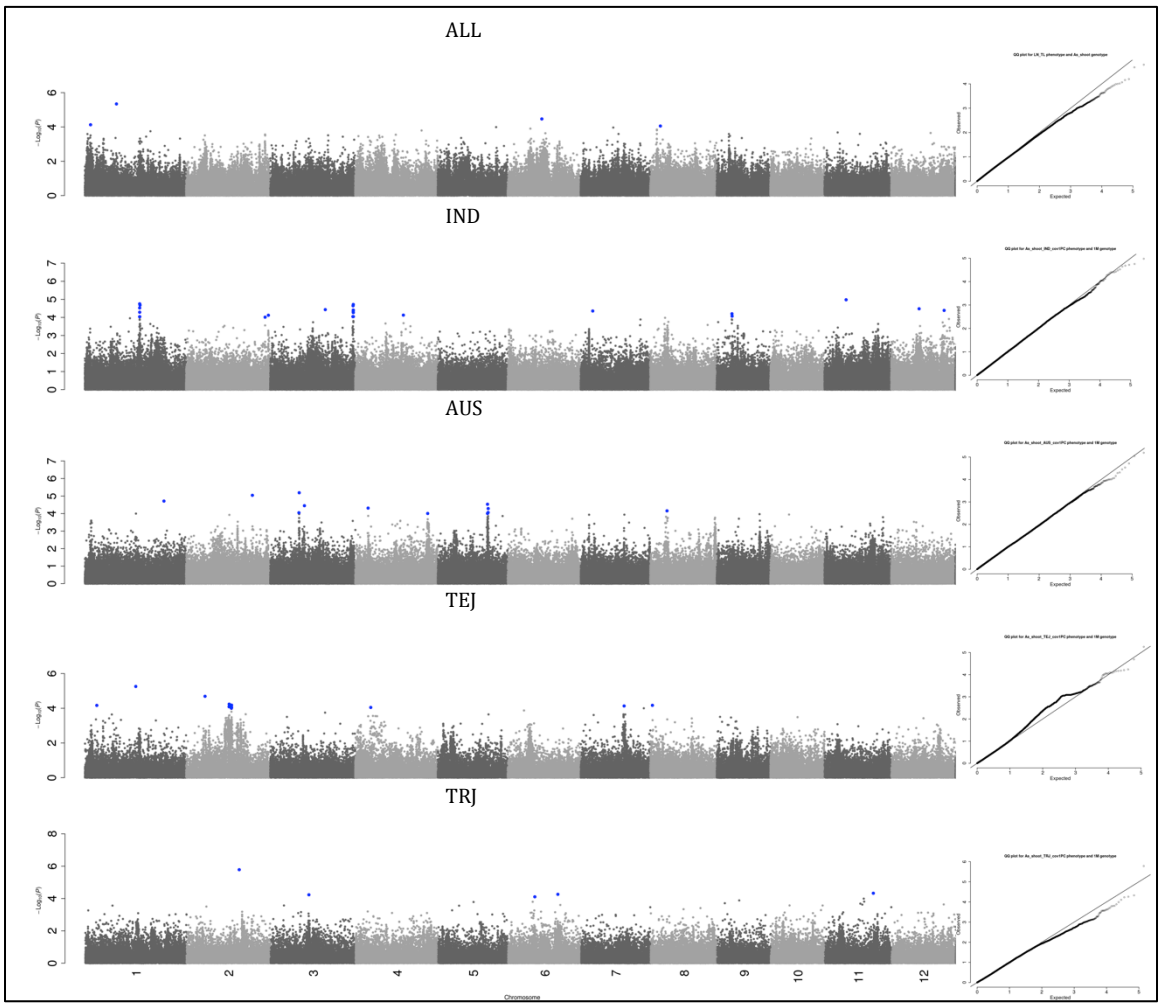
The biplot visual is also a convenient tool to empower comparative ionomics between populations and species. (Chen et al. 2009) considered whole shoot ionomic variation among 45 natural accessions of *L. japonicas* grown hydroponically with subtoxic levels of many heavy metals and metalloids and as such their results lend themselves well to direct comparison with our own. They found As, Fe, Na, Zn, and Mo to be positively correlated in the *L. japonicas* shoot ionome. Our results indicate likewise that As, Fe, Na, and Zn are positively correlated to varying degrees (see clusters (IND(II), IND(III), IND(IV), AUS(IV), AUS(V), AUS(VI), TRJ(IV), TRJ(VI), TEJ(III), and TEJ(IV)) but we found elements almost never share a significant relationship with shoot Mo content. Furthermore, a previous study (Shaibur et al. 2006), demonstrated that As toxicity has been known to induce Fe deficiency in rice shoots, which led to the hypothesis that perhaps a negative correlative relationship should exist between these two phenotypes. However, the analysis of (Shaibur et al. 2006) not only looked at much higher As levels than what was present in our study, but their conclusions were drawn from a single *temperate japonica* variety. Our data shows that As and Fe are positively correlated to varying degrees across subpopulations and that the correlation is the weakest in the *temperate japonica* gene

pool (see clusters TEJ(III) and TEJ(IV)).

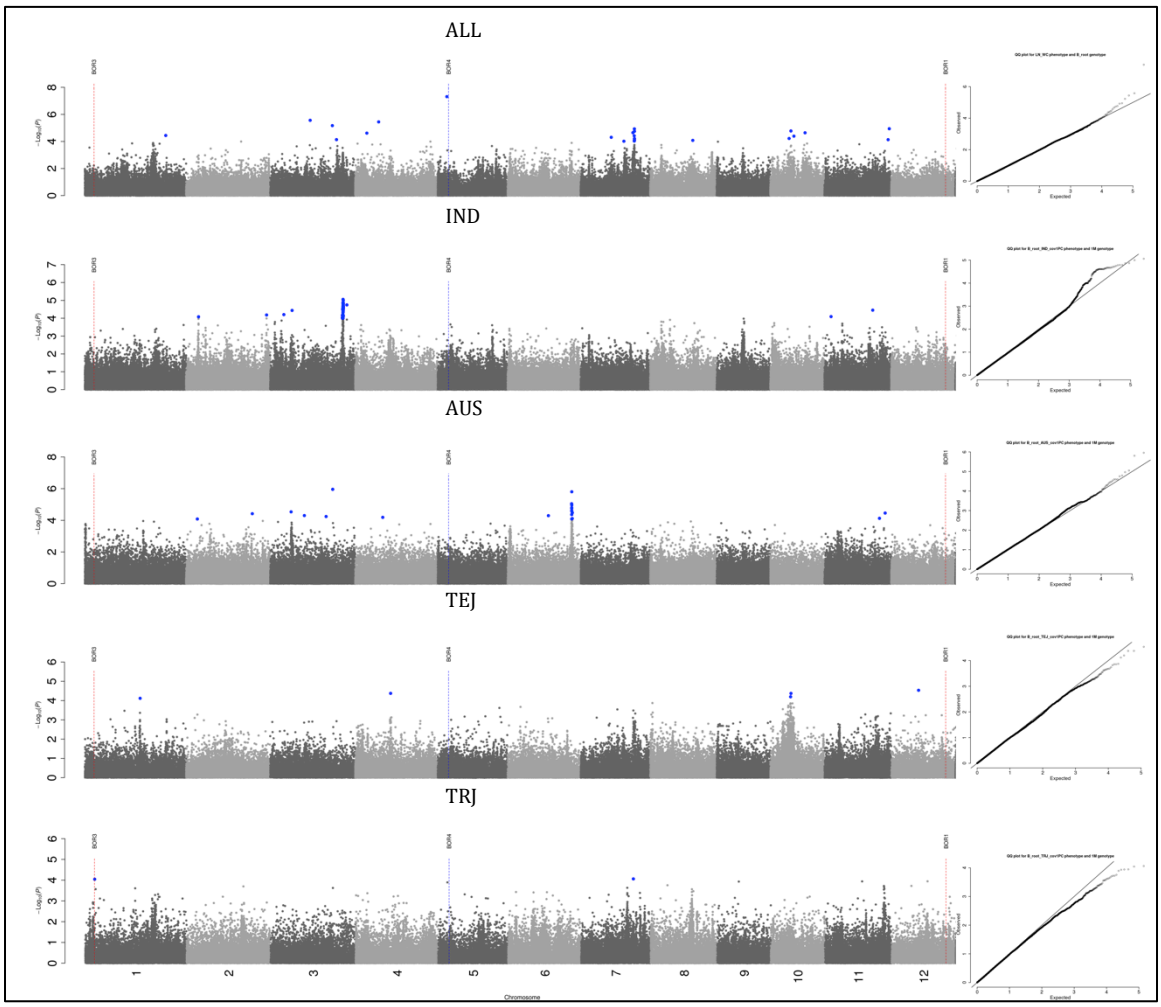
In *Arabidopsis thaliana*, (Lahner et al. 2003) found that Fe accumulation in *A. thaliana* shoots affected phosphate content, but our analysis showed a very weak relationship between Fe and P concentrations in rice shoots. It has however, been suggested that potassium is involved in the increased absorption and distribution of Fe in plants (Alam et al. 2002, Jolley et al. 1988) and that deficiencies in either one of these elements impacts the expression profiles of many of the same gene networks (Wang et al. 2002). Interestingly, in the present analysis the *Indica* clade shows slightly positive correlations between Fe and K while the *Japonica* clade harbors slightly negative correlations. This could perhaps be an indication that differing molecular mechanisms between the two divergent clades are influencing the crosstalk between K and Fe pathways.



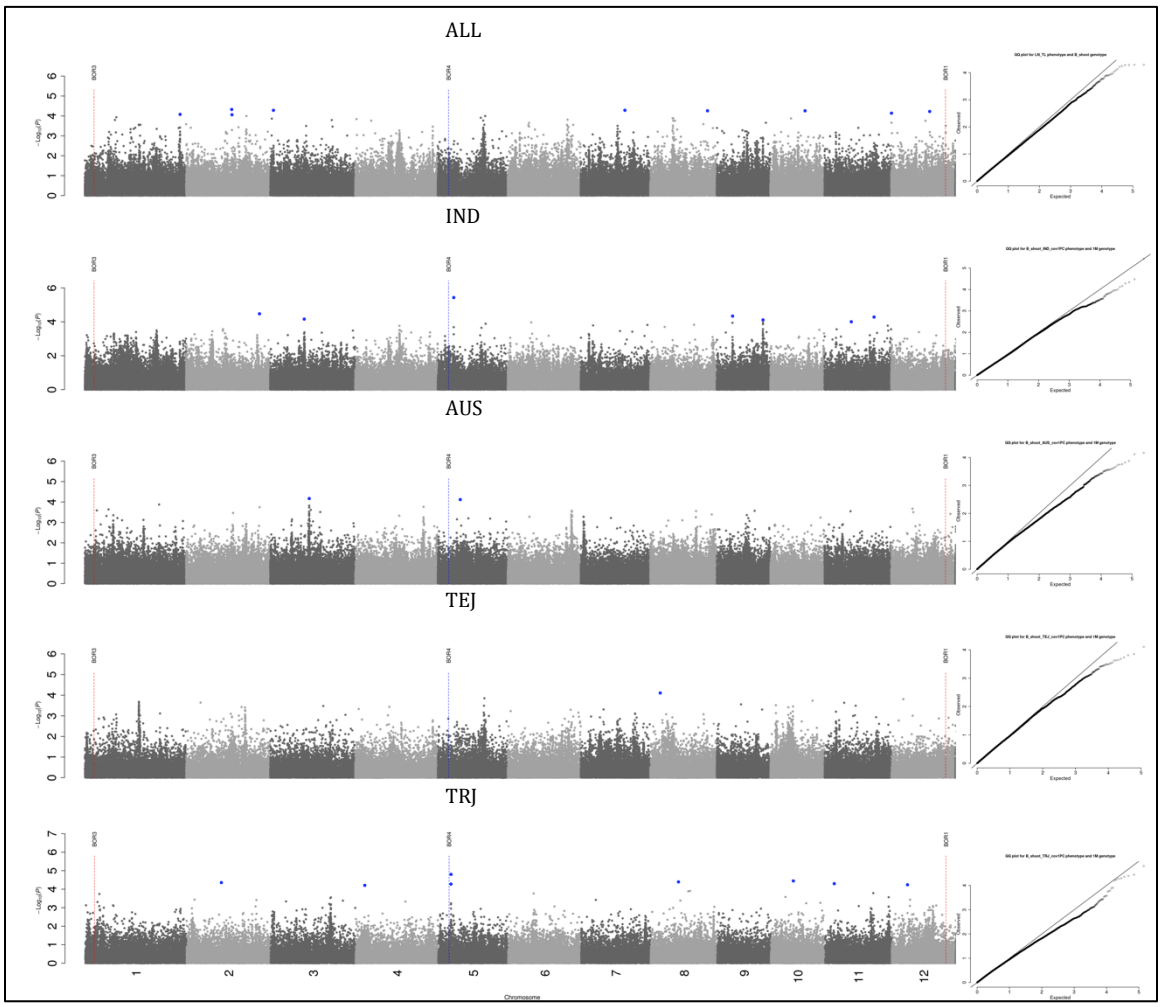
Supplementary Figure 3.3 Manhattan and quantile-quantile plots for arsenic (As) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



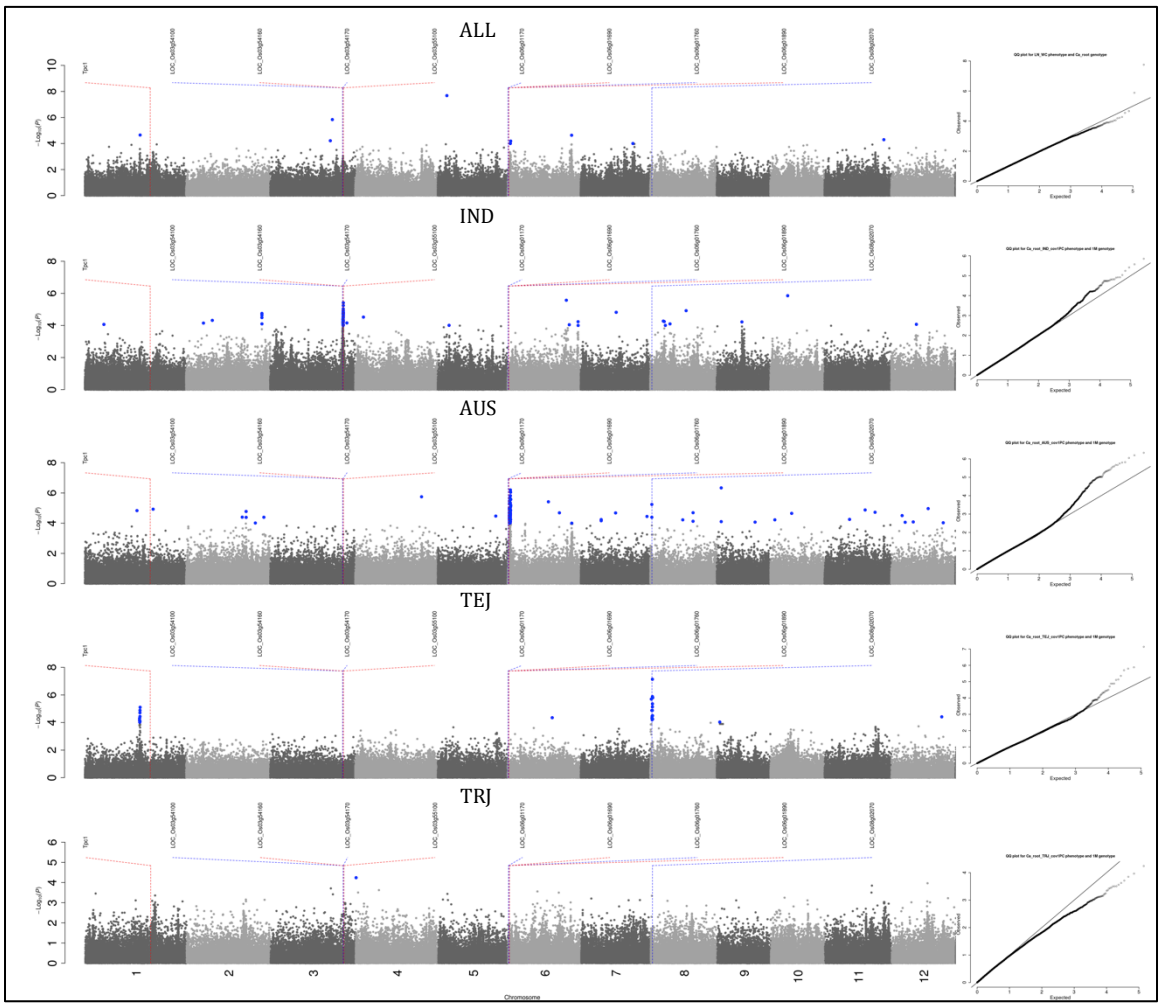
Supplementary Figure 3.4 Manhattan and quantile-quantile plots for arsenic (As) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



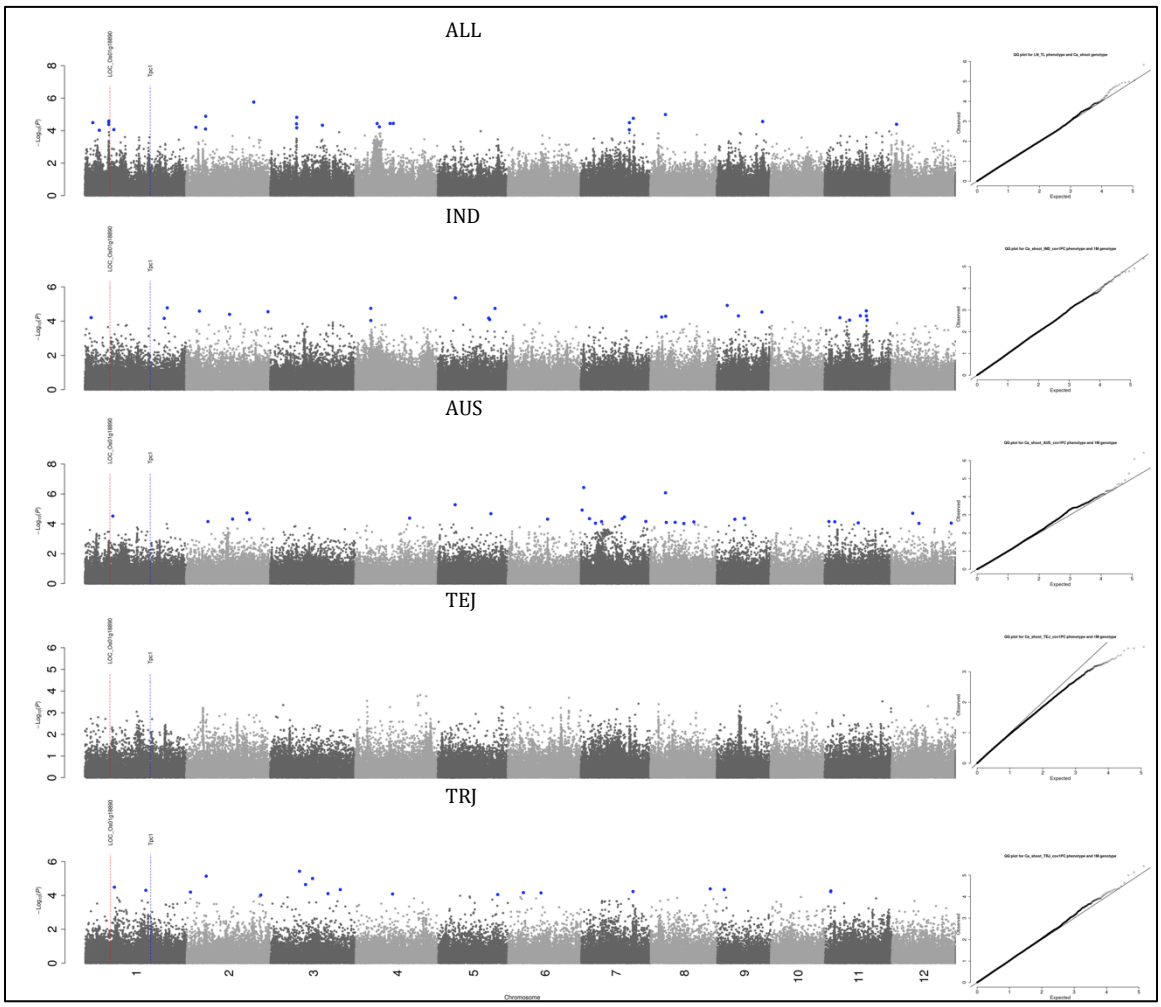
Supplementary Figure 3.5 Manhattan and quantile-quantile plots for boron (B) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



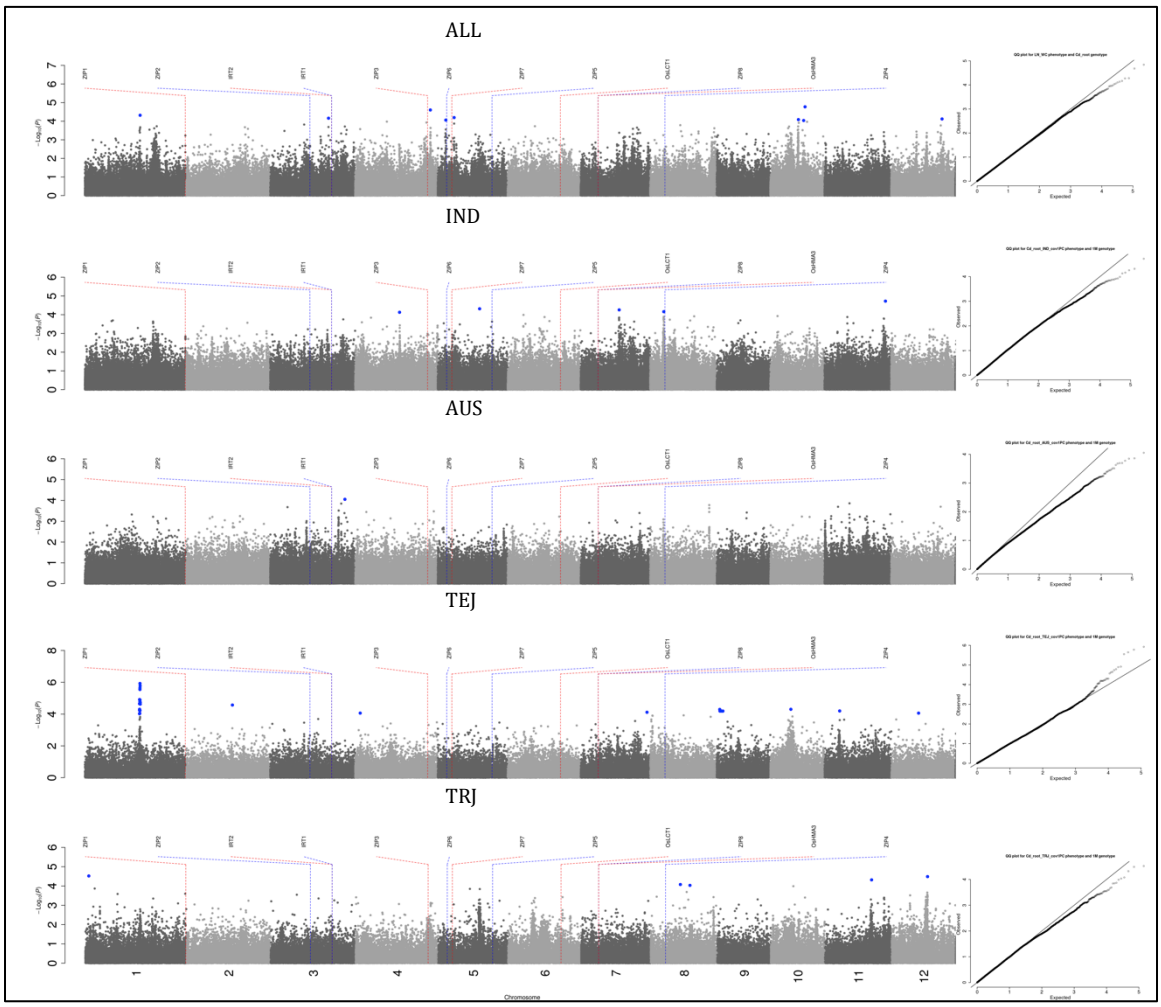
Supplementary Figure 3.6 Manhattan and quantile-quantile plots for boron (B) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



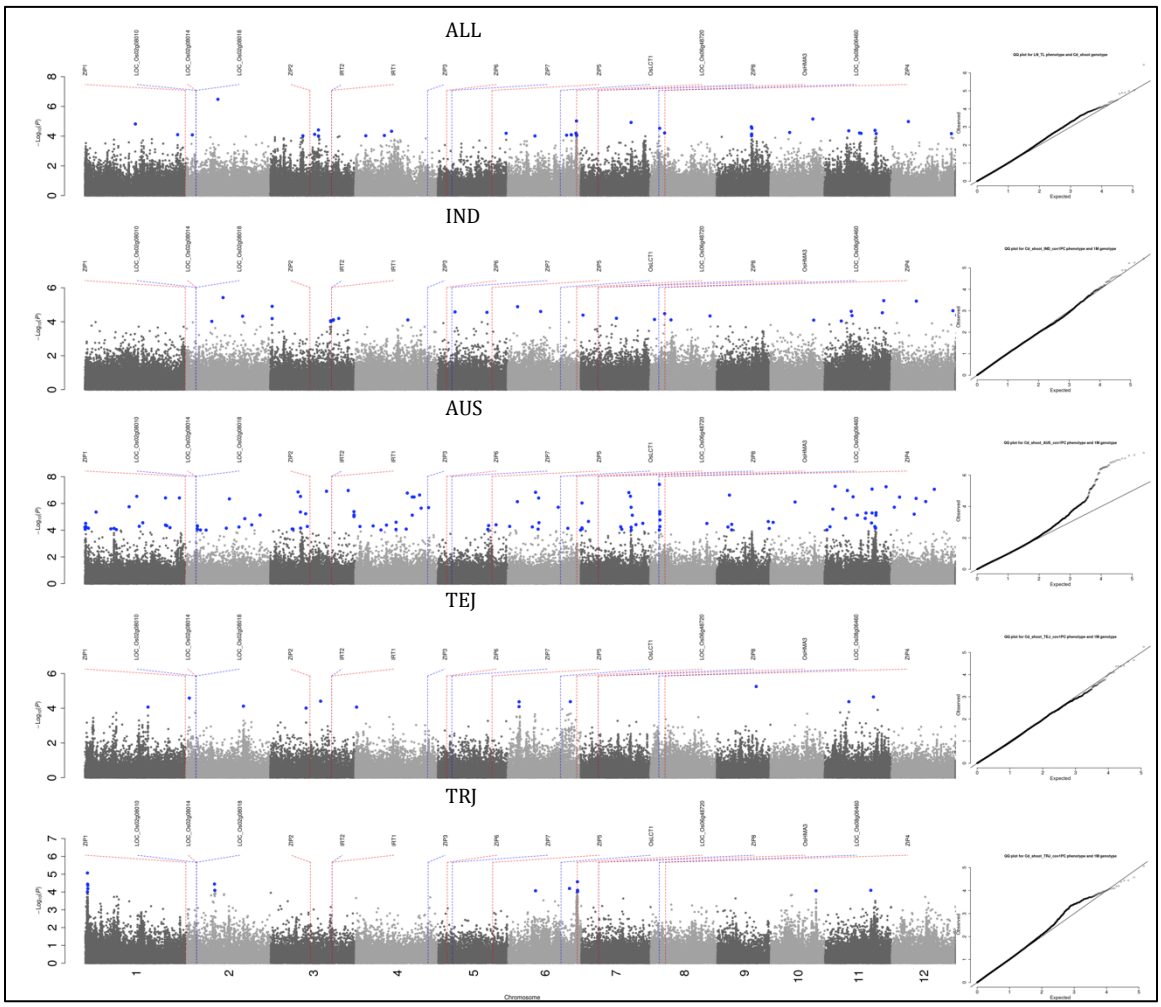
Supplementary Figure 3.7 Manhattan and quantile-quantile plots for calcium (Ca) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



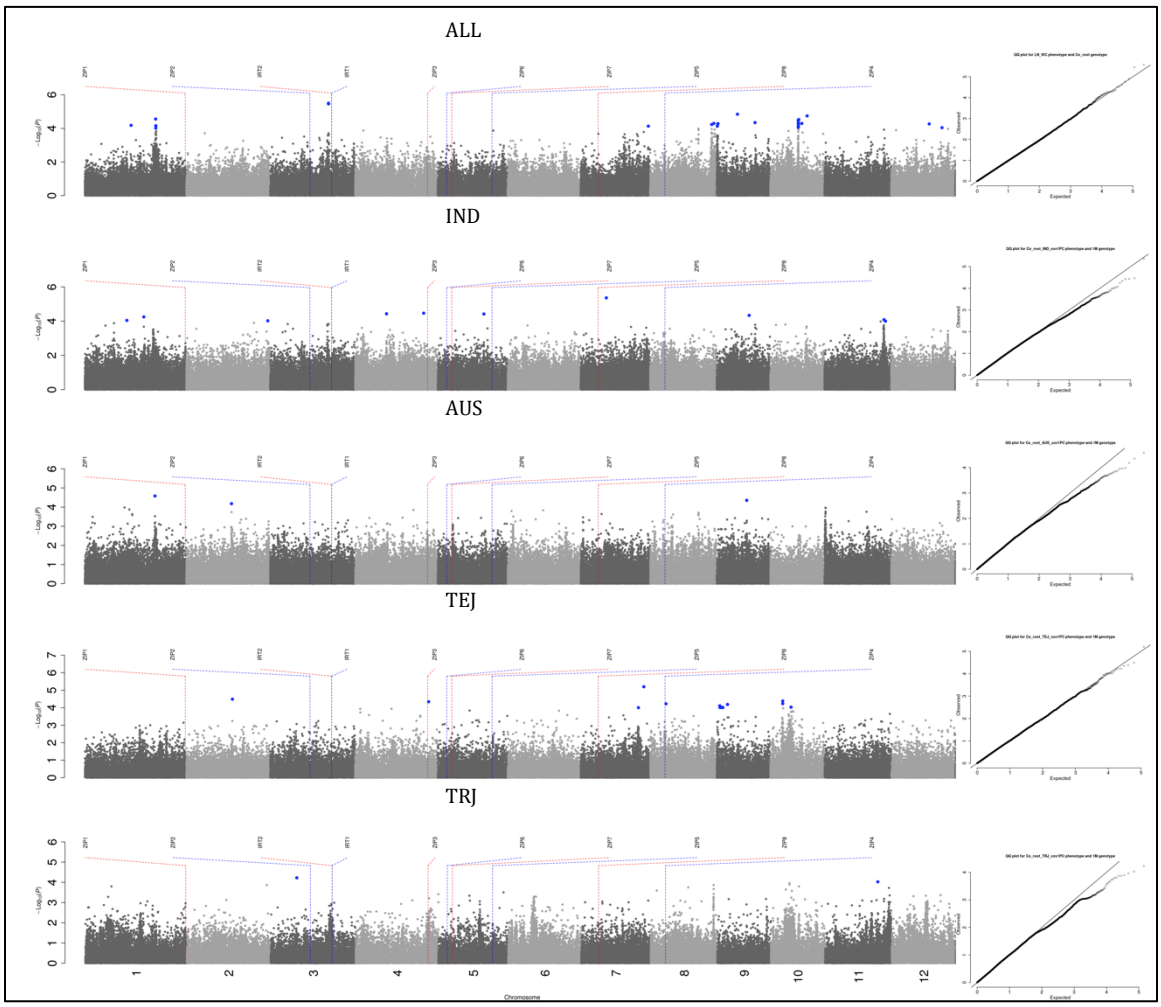
Supplementary Figure 3.8 Manhattan and quantile-quantile plots for calcium (Ca) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



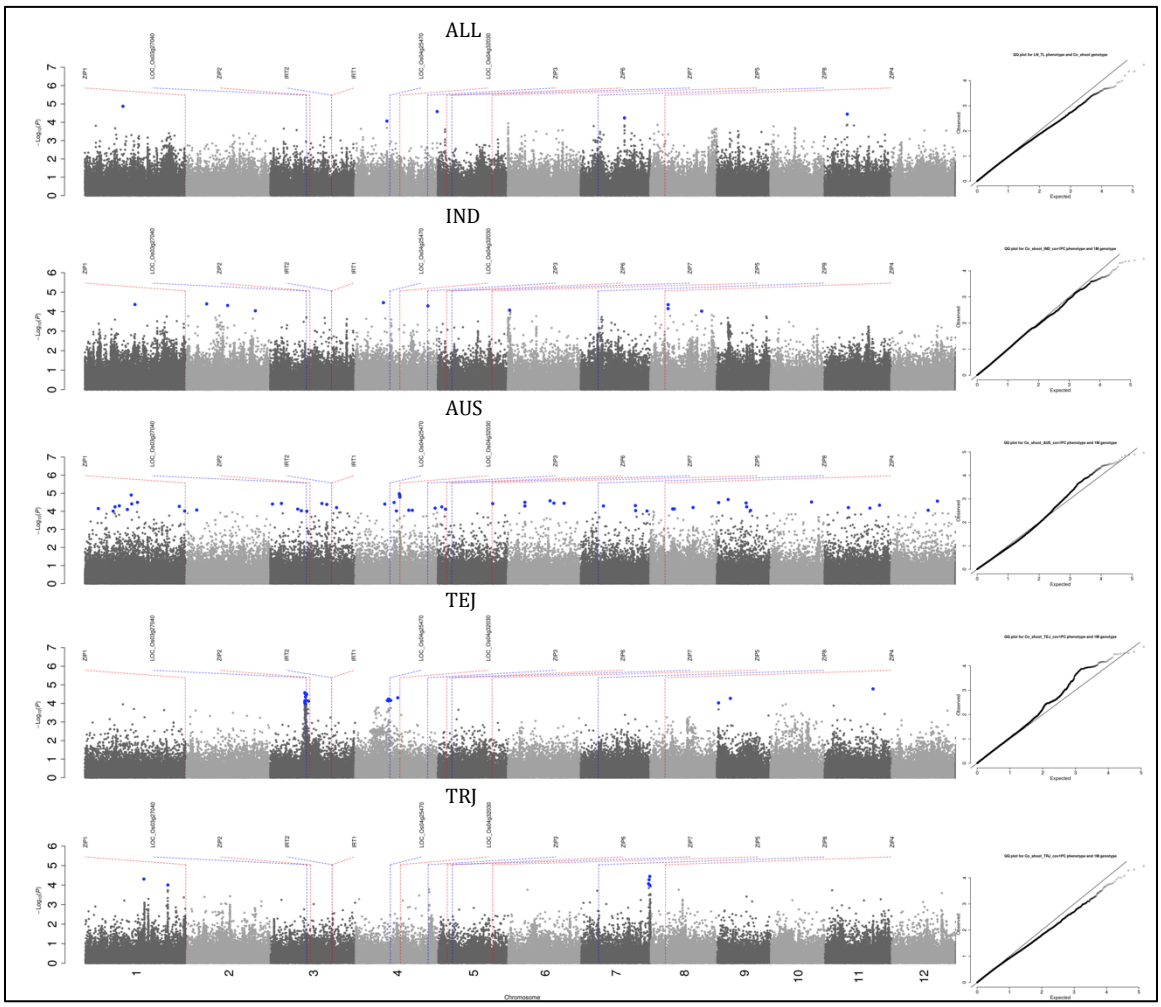
Supplementary Figure 3.9 Manhattan and quantile-quantile plots for cadmium (Cd) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



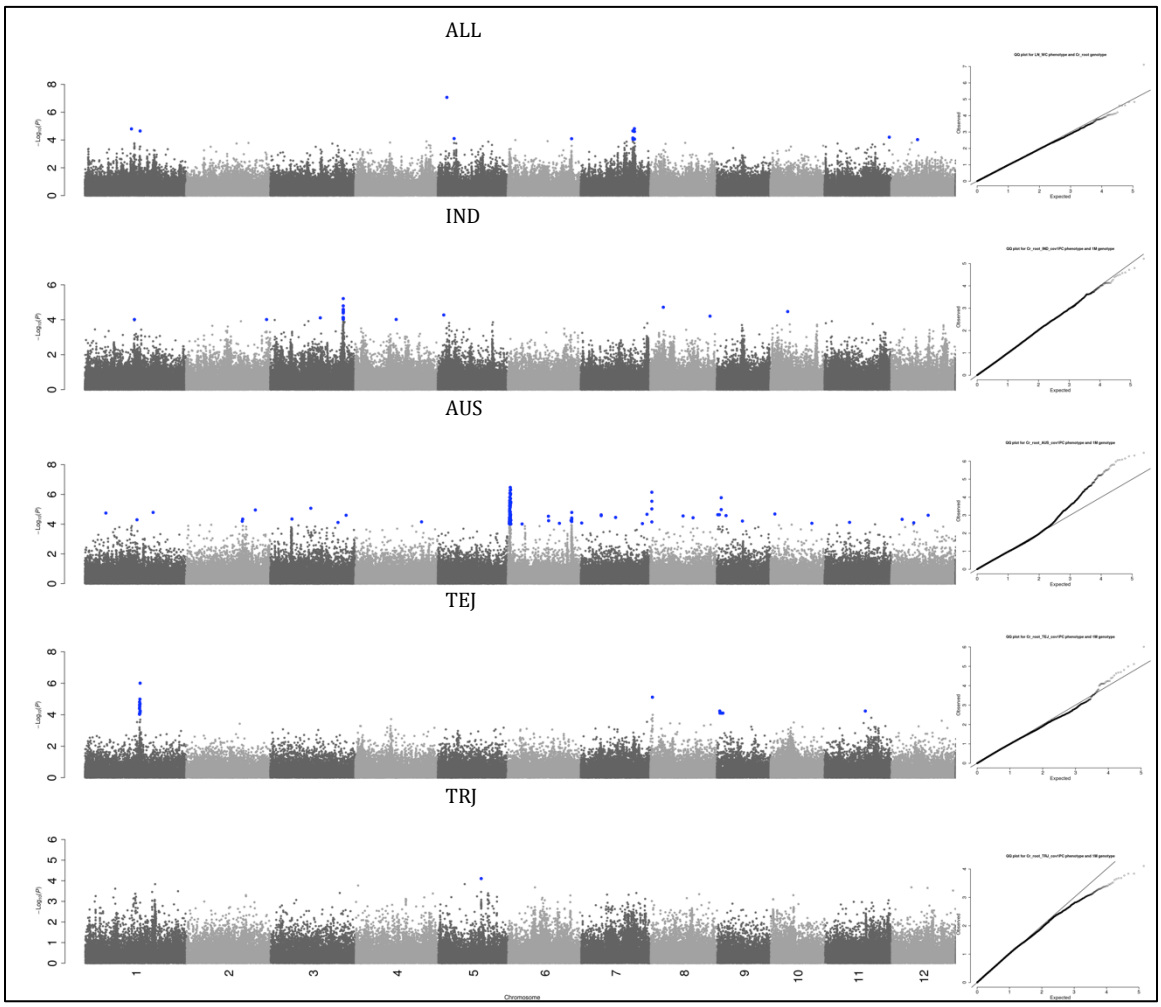
Supplementary Figure 3.10 Manhattan and quantile-quantile plots for cadmium (Cd) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



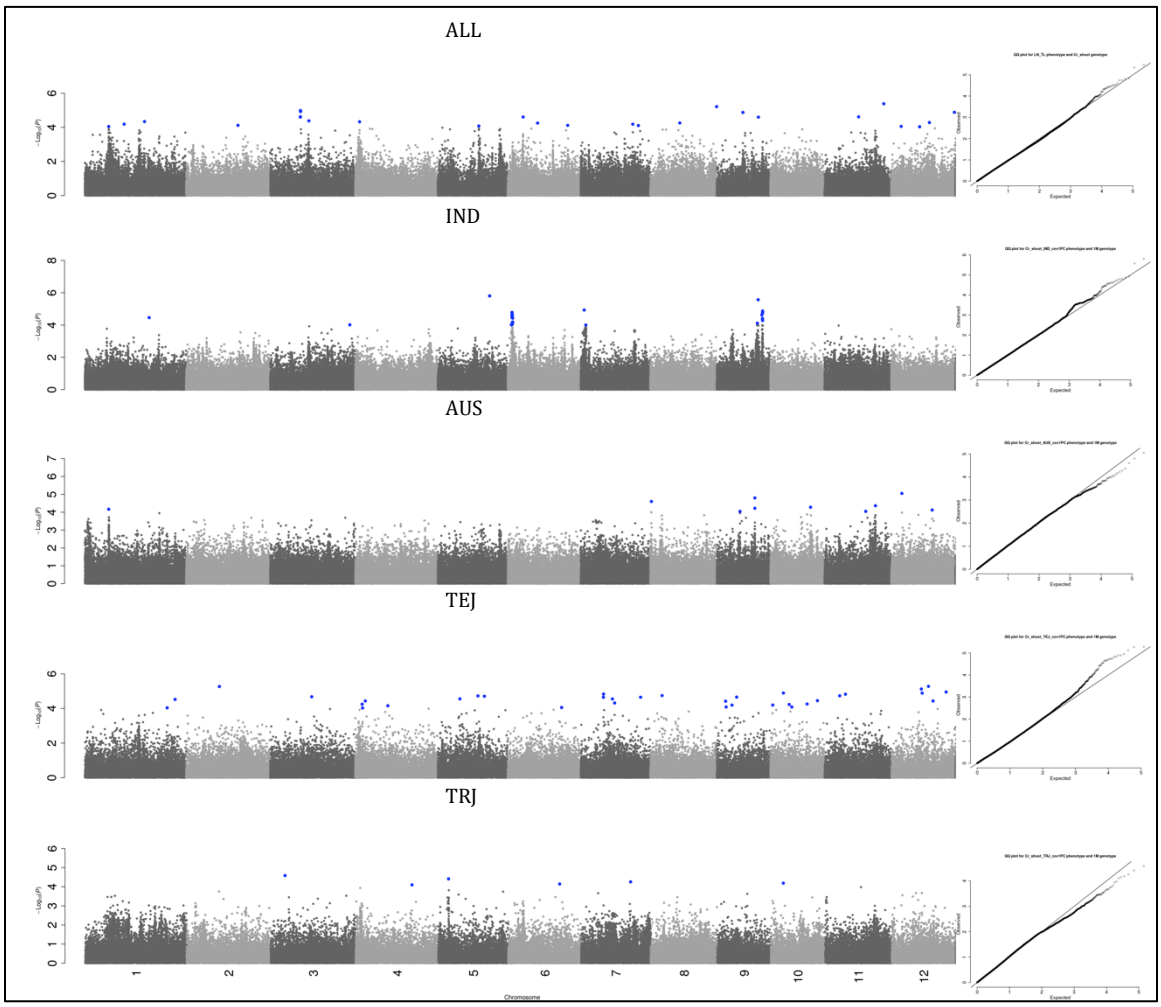
Supplementary Figure 3.11 Manhattan and quantile-quantile plots for cobalt (Co) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



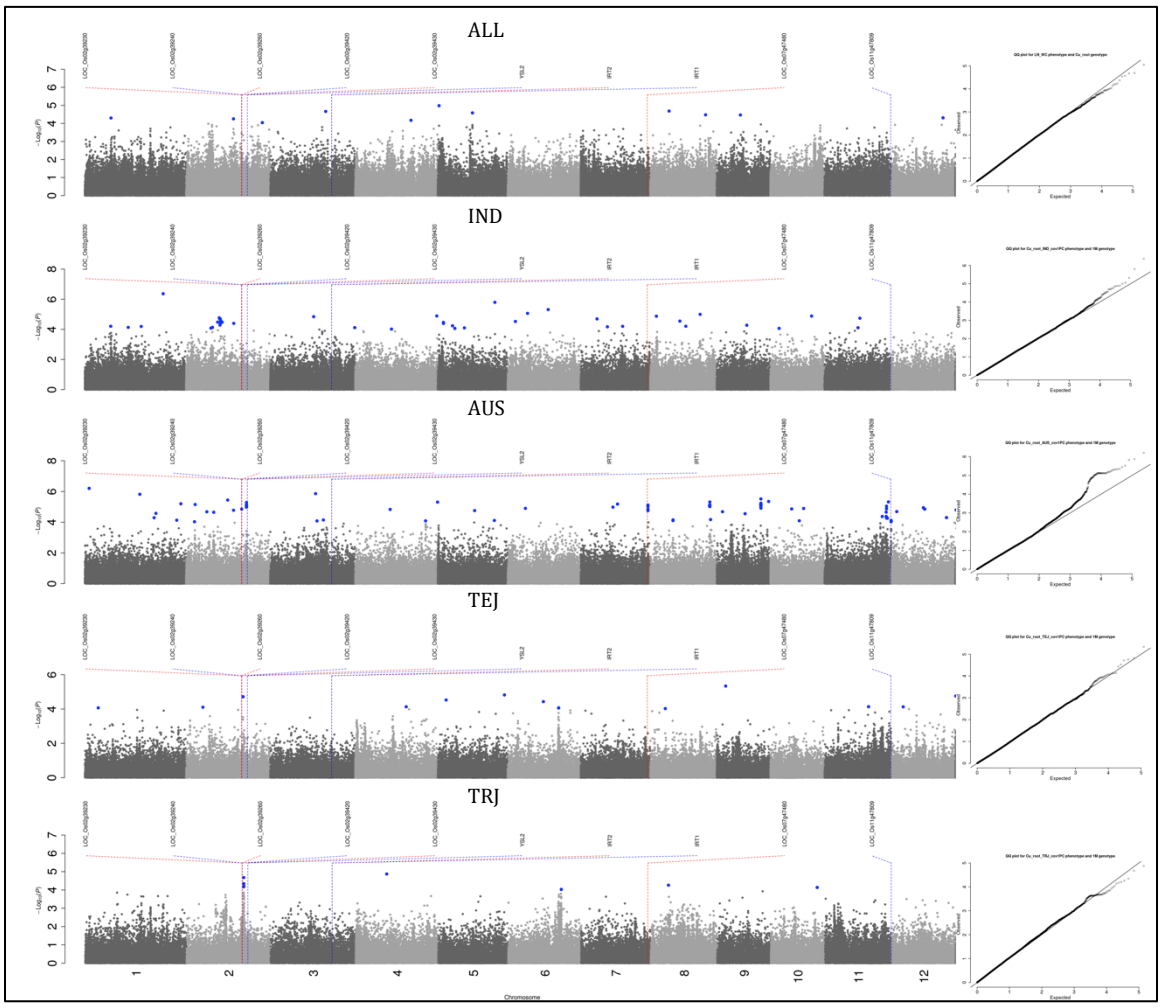
Supplementary Figure 3.12 Manhattan and quantile-quantile plots for cobalt (Co) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



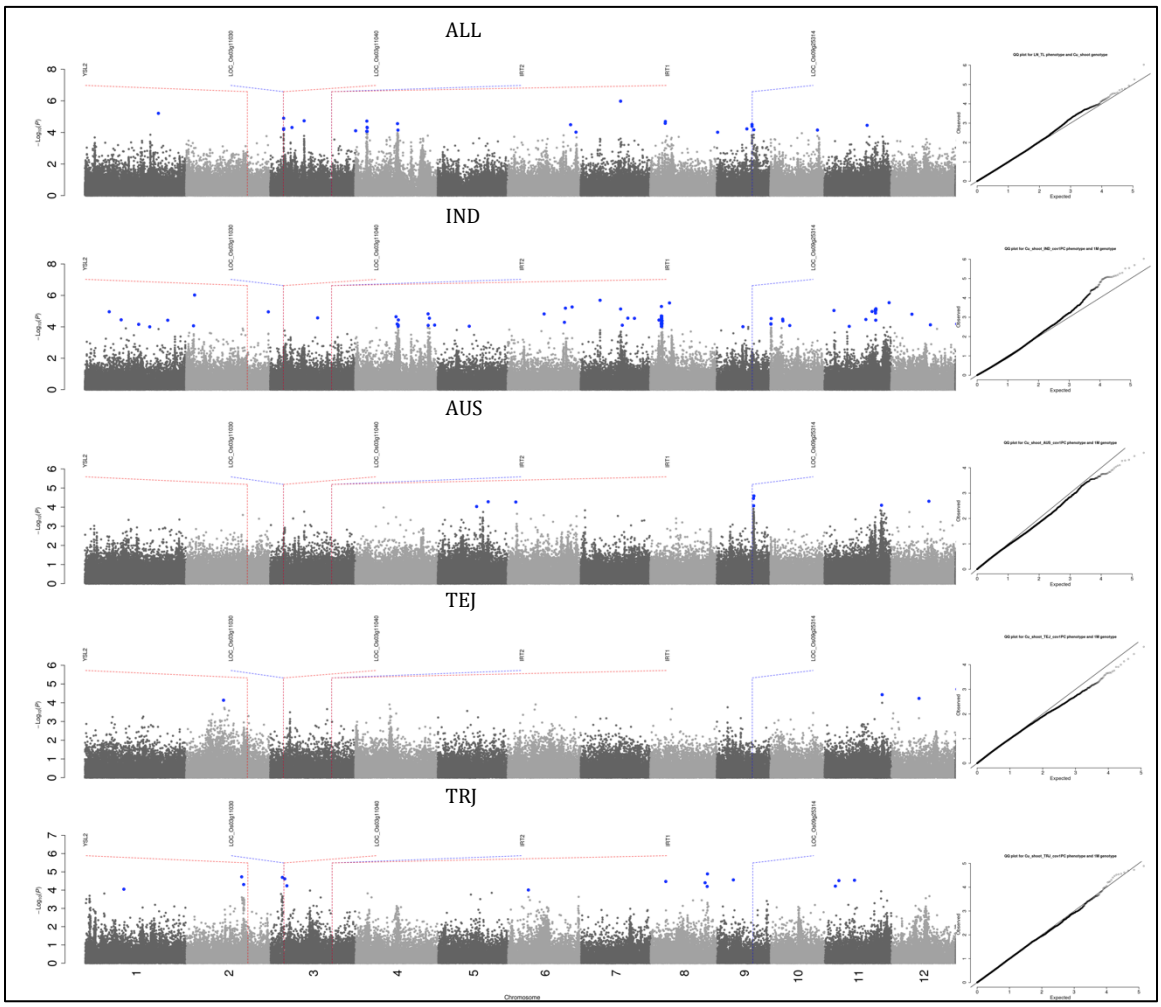
Supplementary Figure 3.13 Manhattan and quantile-quantile plots for chromium (Cr) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



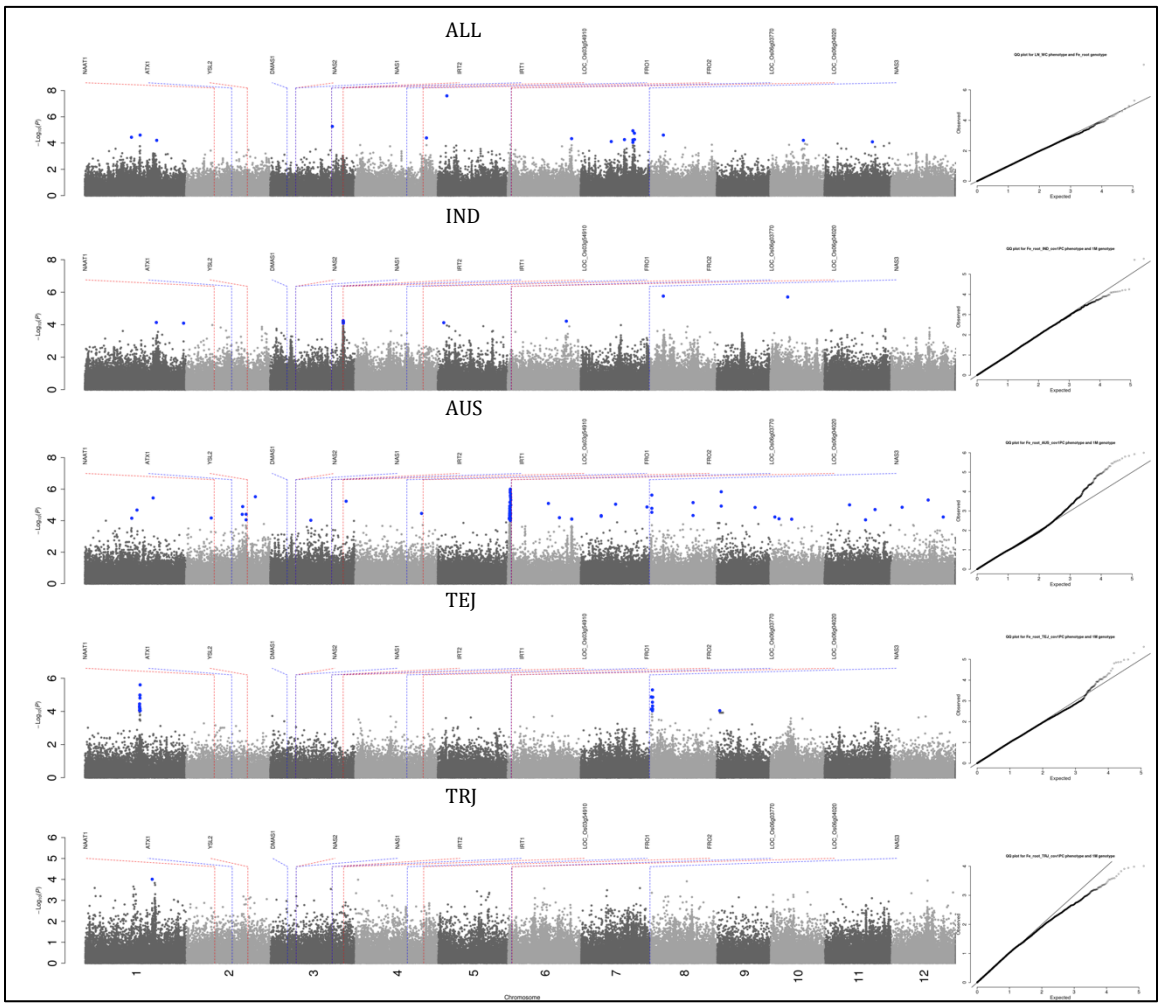
Supplementary Figure 3.14 Manhattan and quantile-quantile plots for chromium (Cr) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



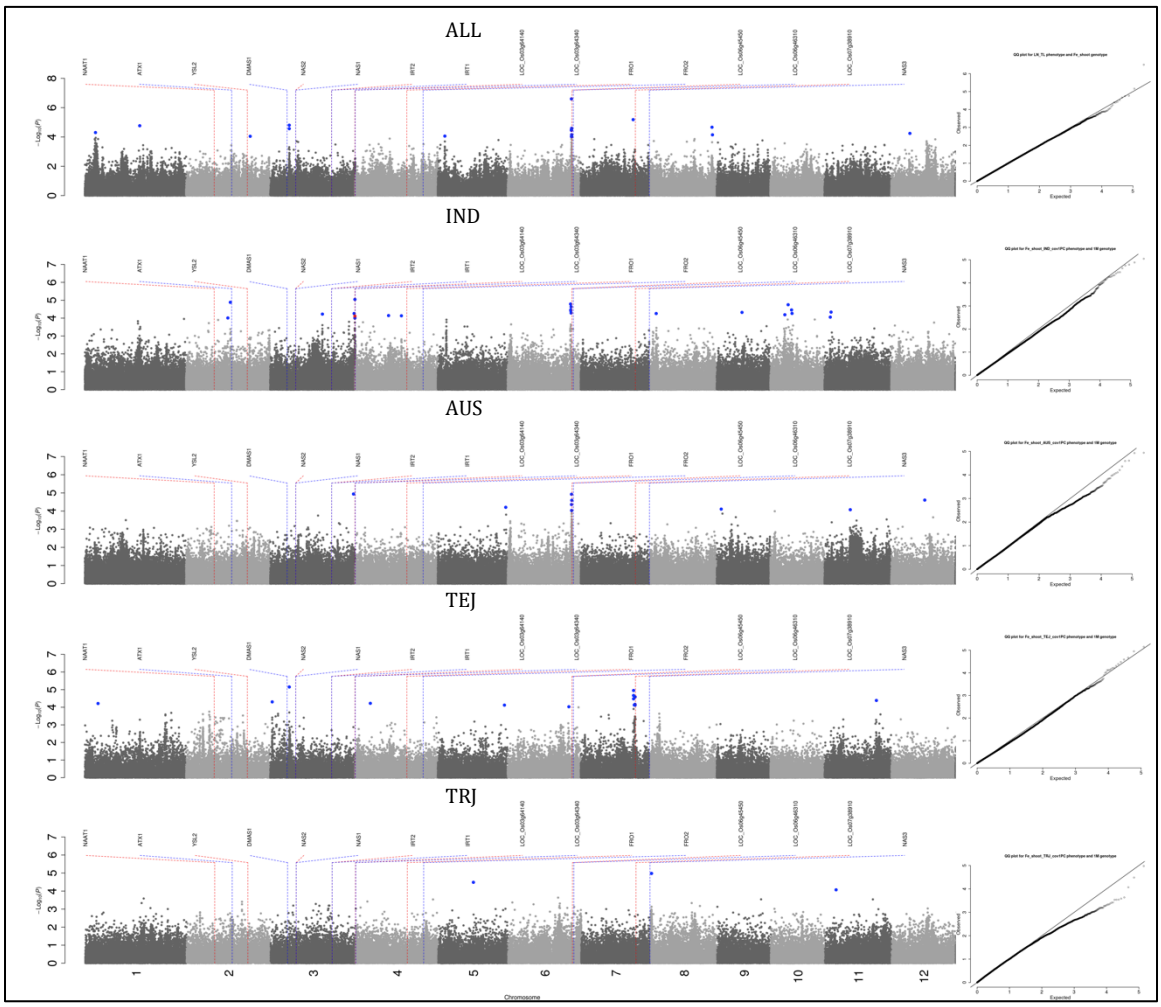
Supplementary Figure 3.15 Manhattan and quantile-quantile plots for copper (Cu) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



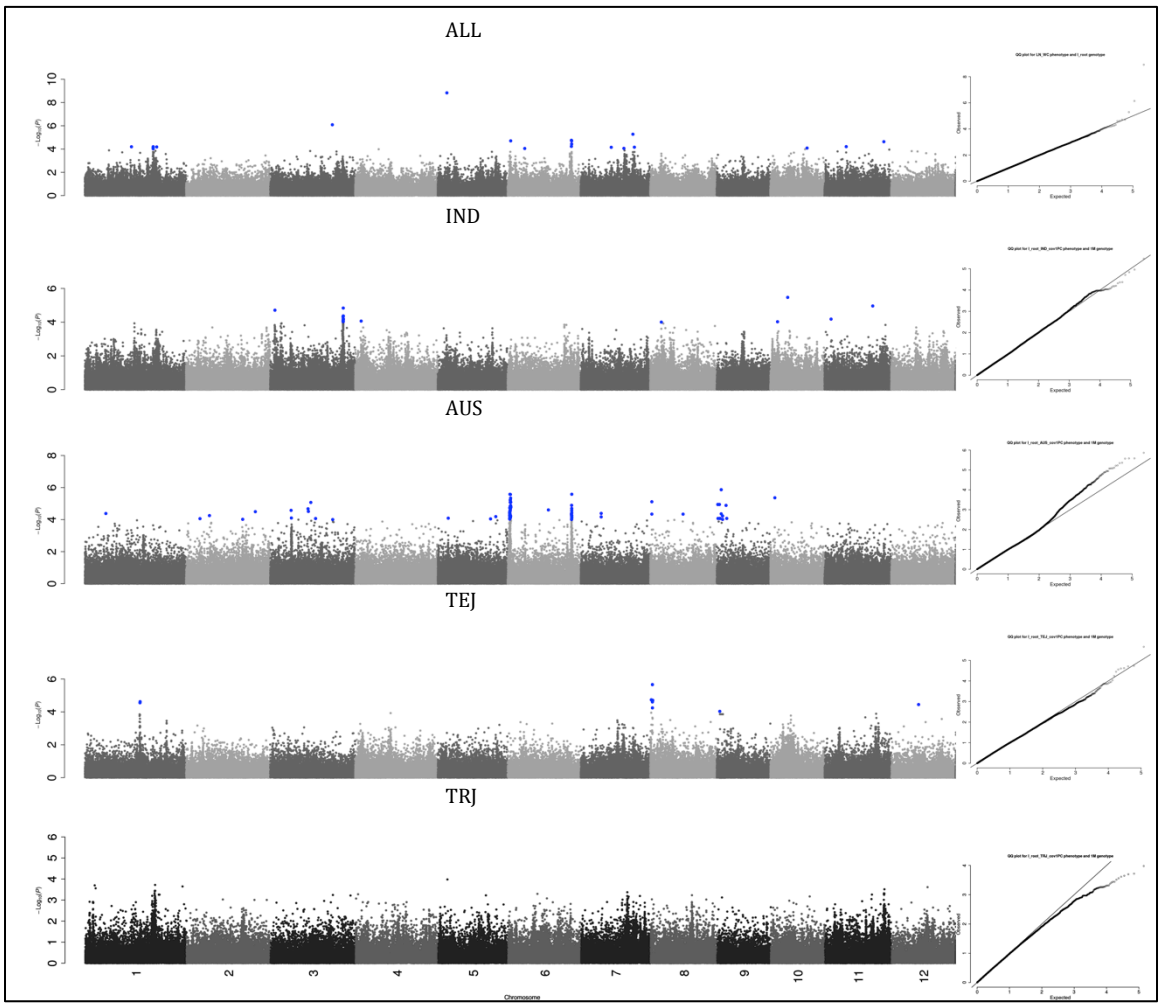
Supplementary Figure 3.16 Manhattan and quantile-quantile plots for copper (Cu) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



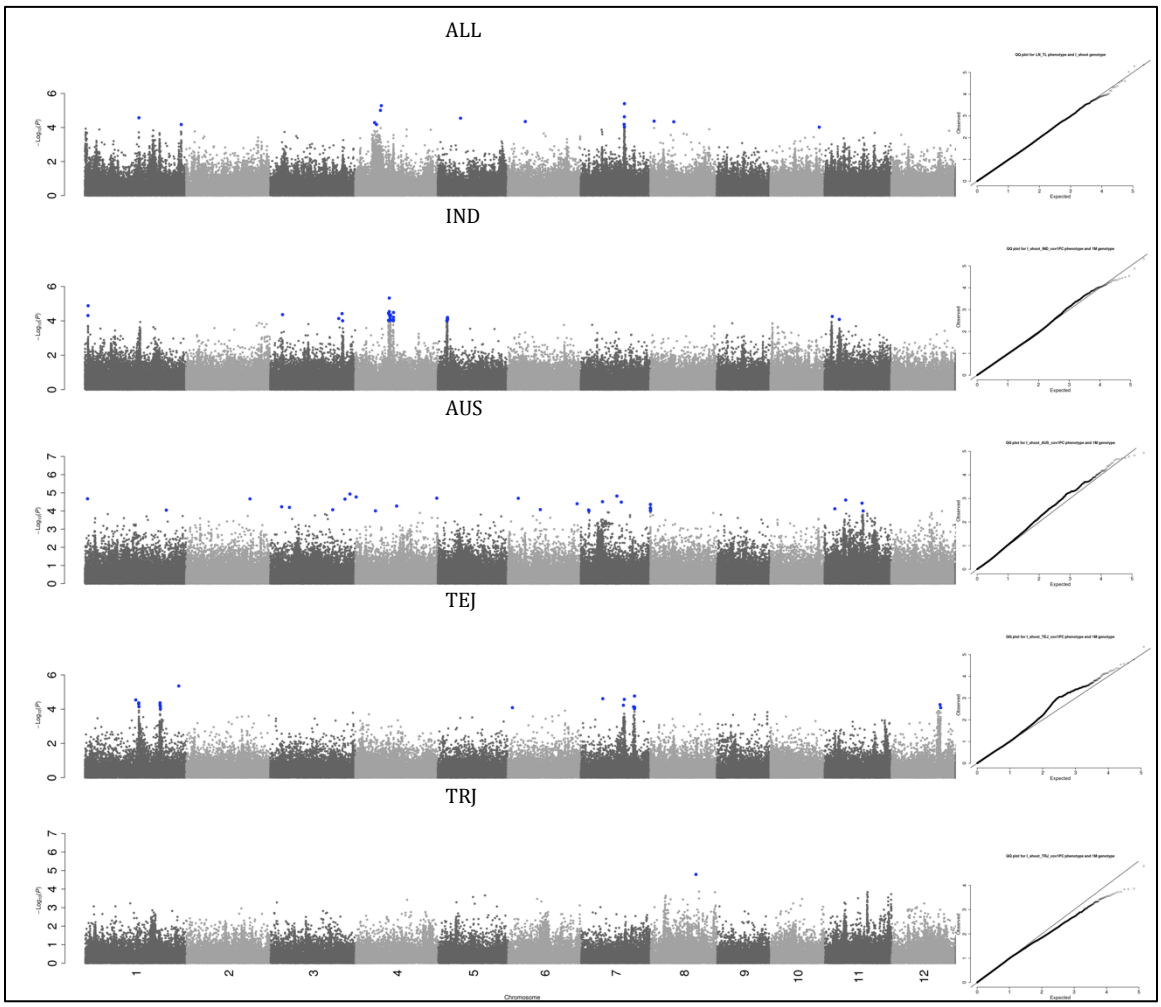
Supplementary Figure 3.17 Manhattan and quantile-quantile plots for iron (Fe) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



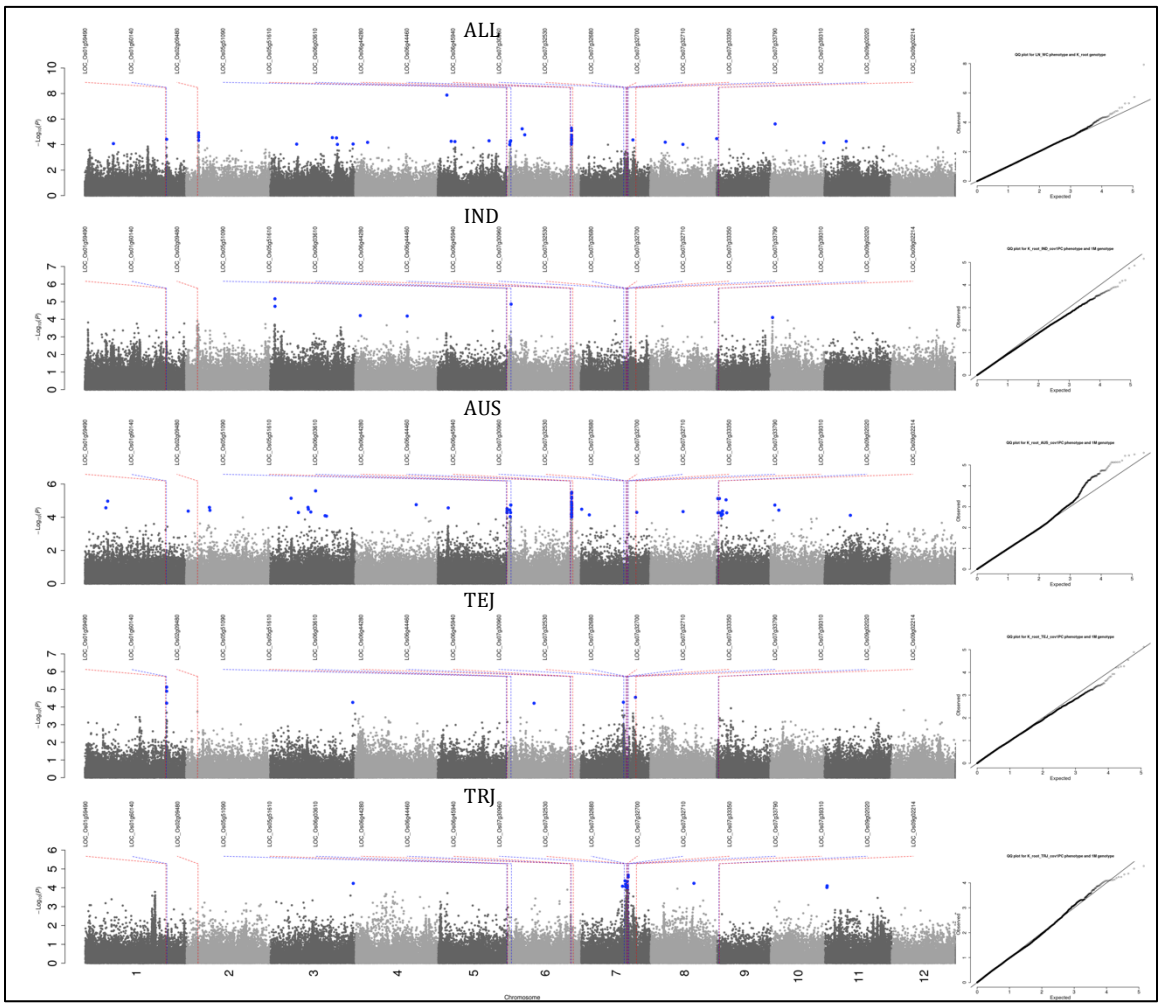
Supplementary Figure 3.18 Manhattan and quantile-quantile plots for iron (Fe) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



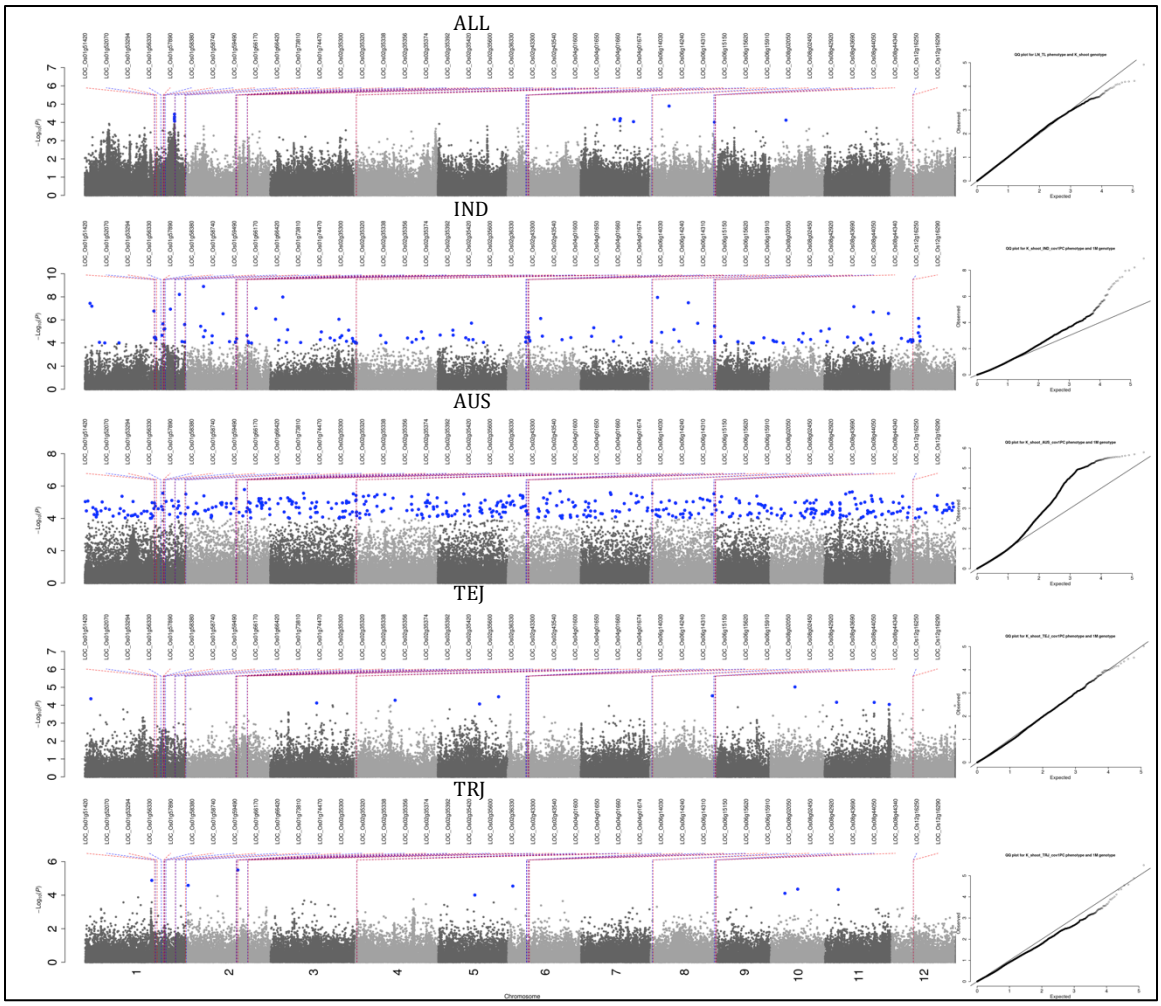
Supplementary Figure 3.19 Manhattan and quantile-quantile plots for iodine (I) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



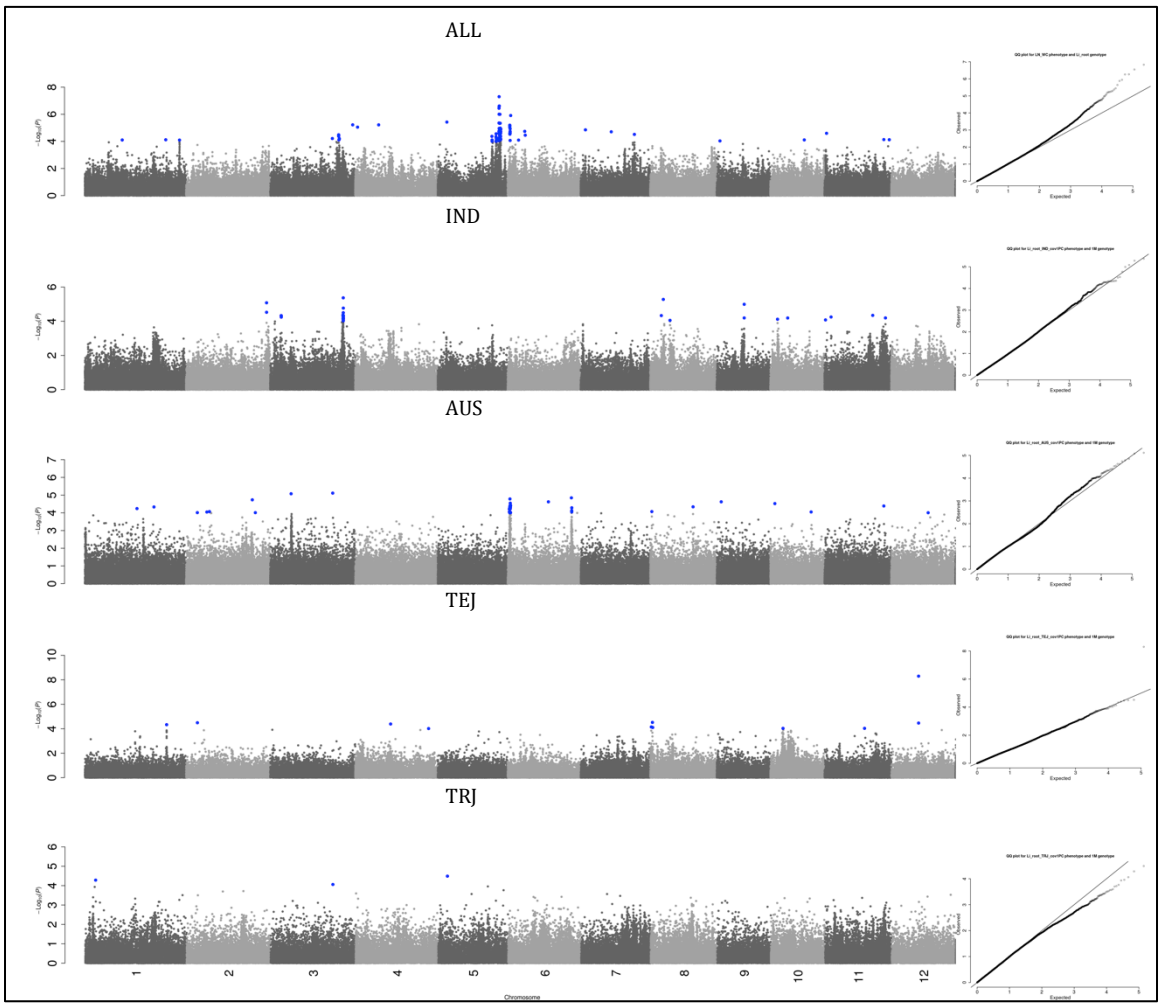
Supplementary Figure 3.20 Manhattan and quantile-quantile plots for iodine (I) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



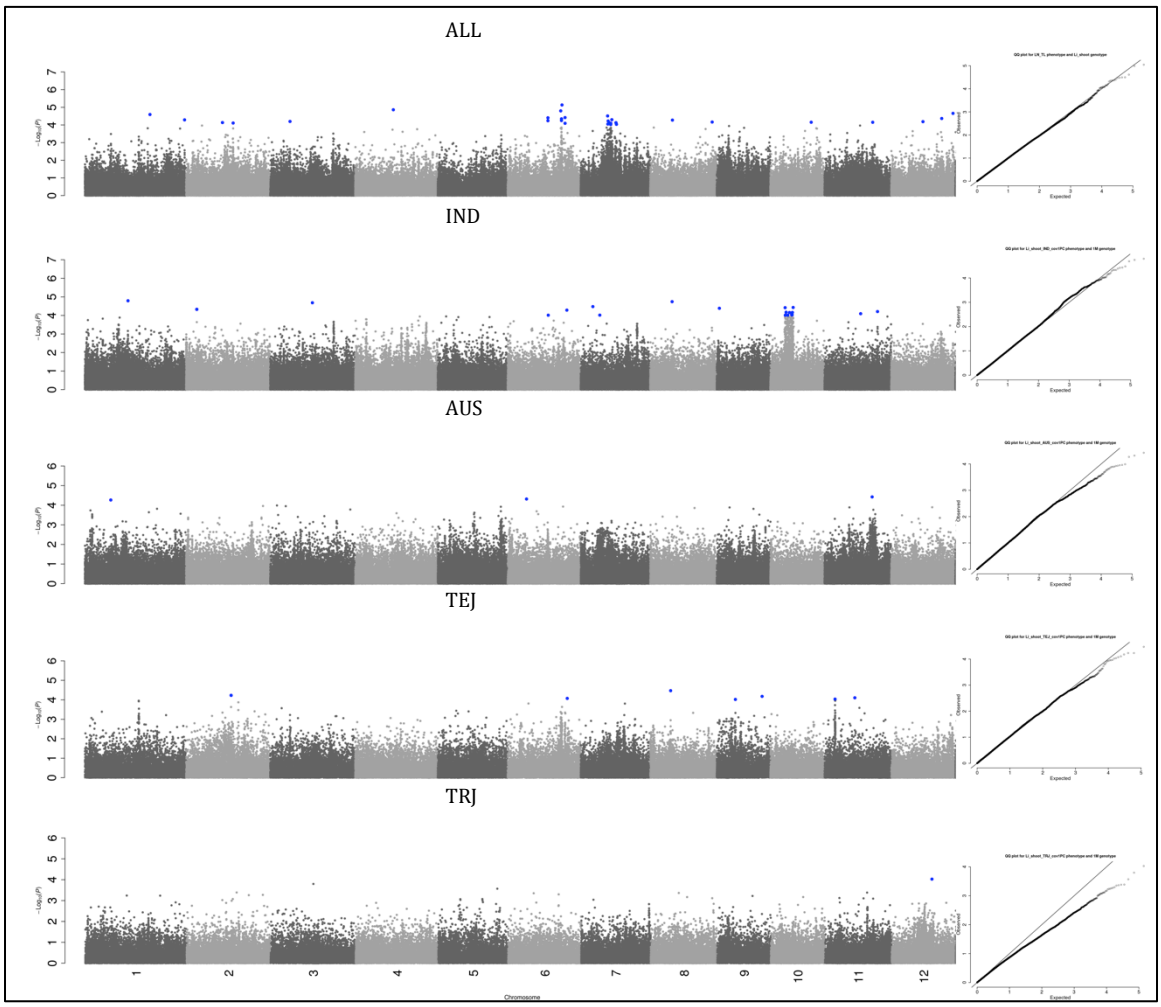
Supplementary Figure 3.21 Manhattan and quantile-quantile plots for potassium (K) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



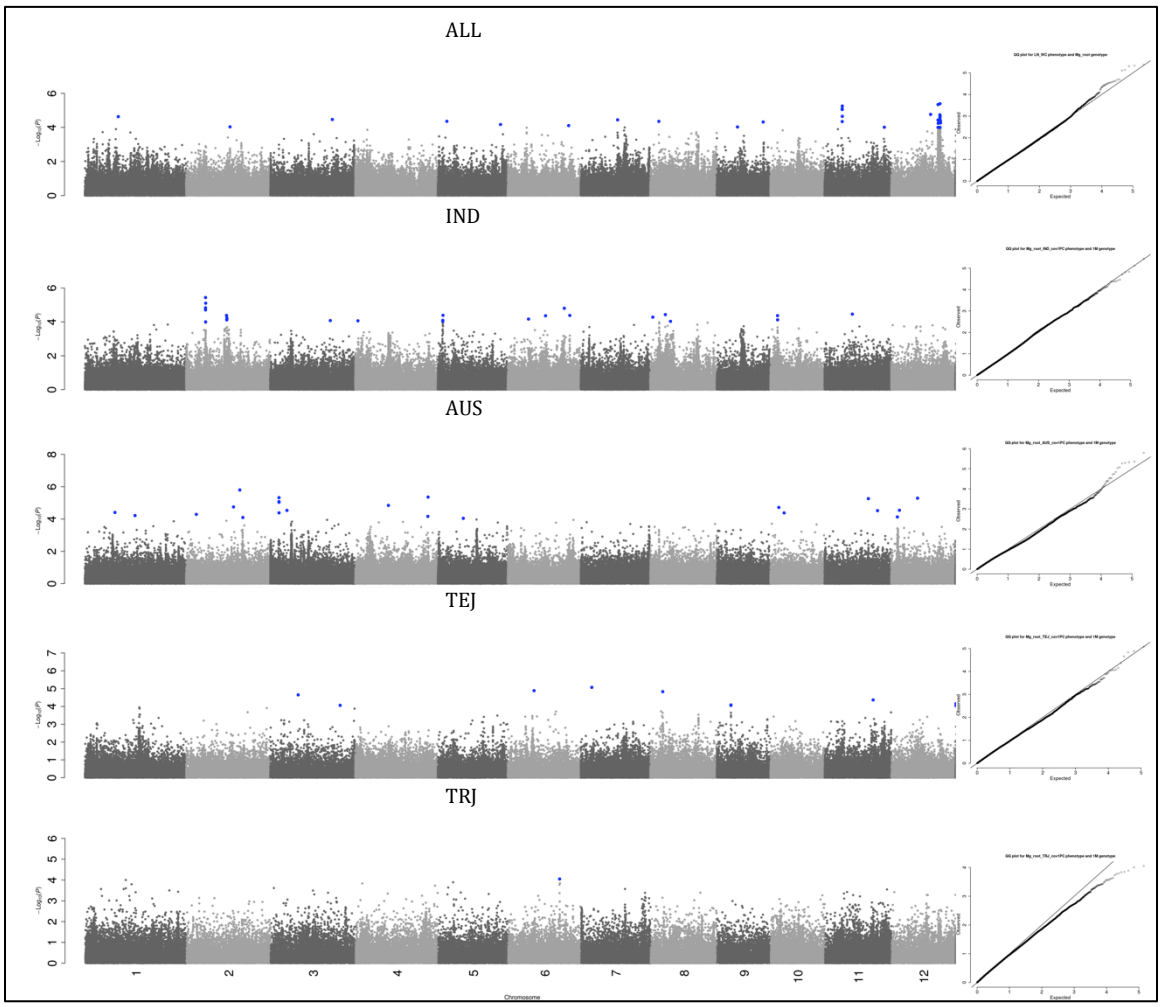
Supplementary Figure 3.22 Manhattan and quantile-quantile plots for potassium (K) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



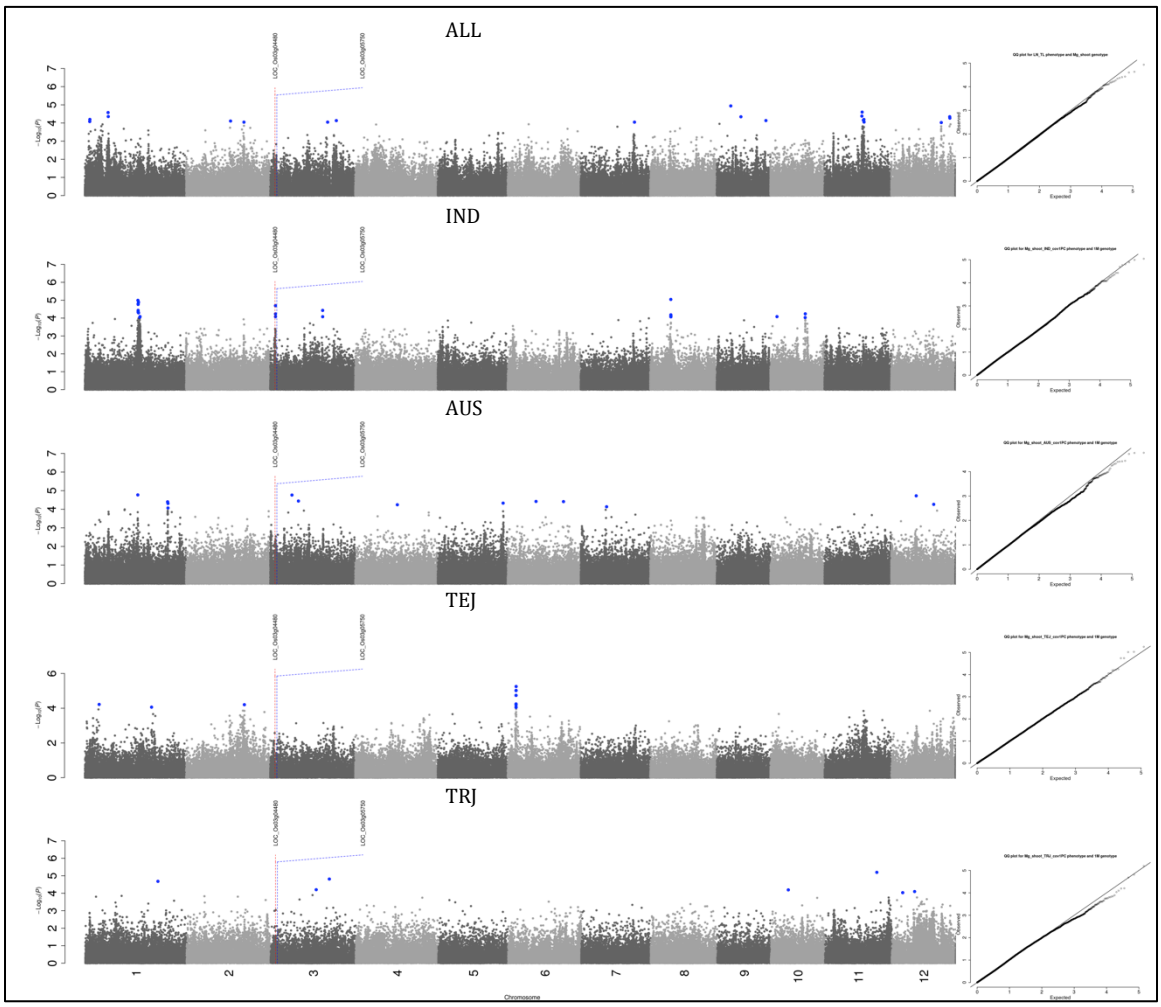
Supplementary Figure 3.23 Manhattan and quantile-quantile plots for lithium (Li) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



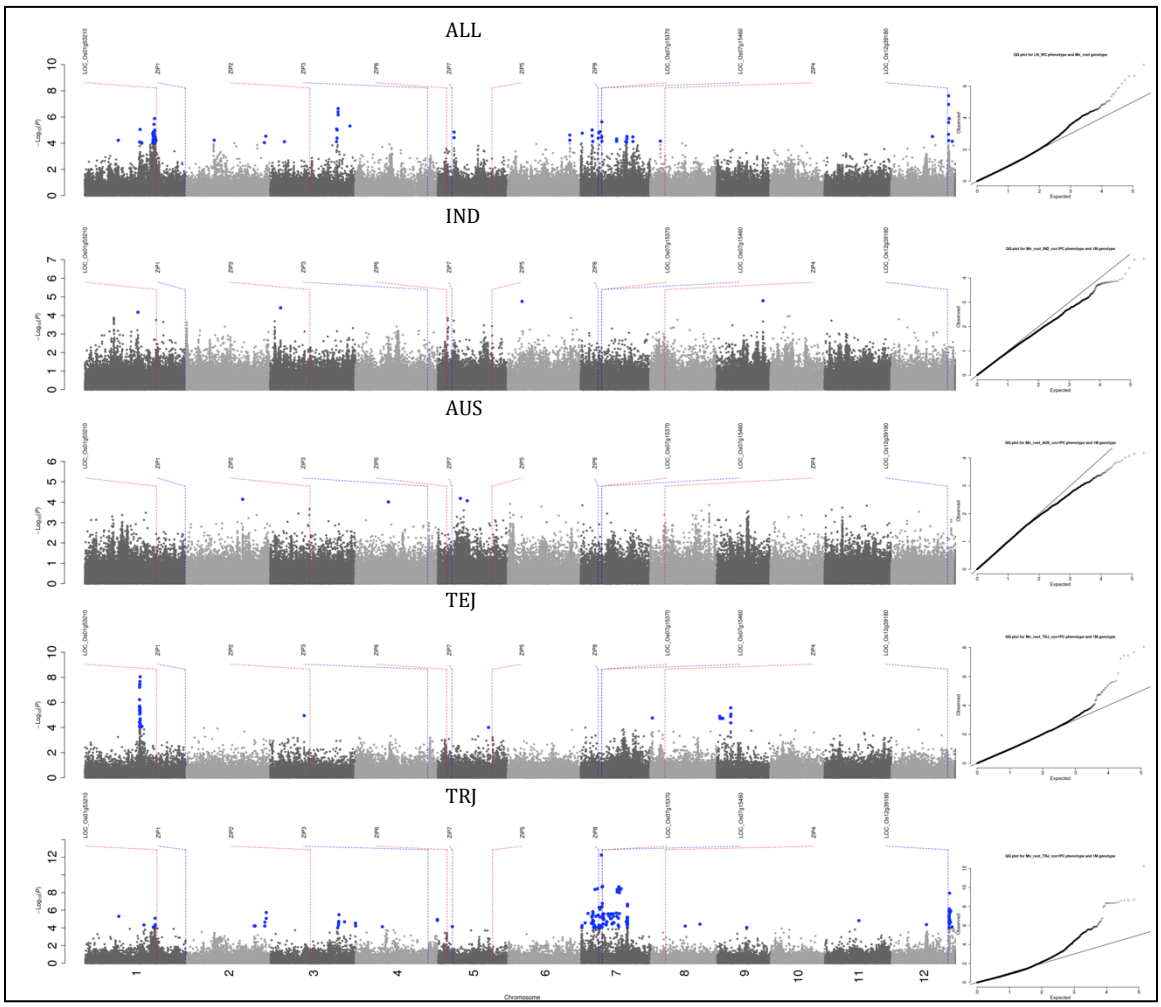
Supplementary Figure 3.24 Manhattan and quantile-quantile plots for lithium (Li) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



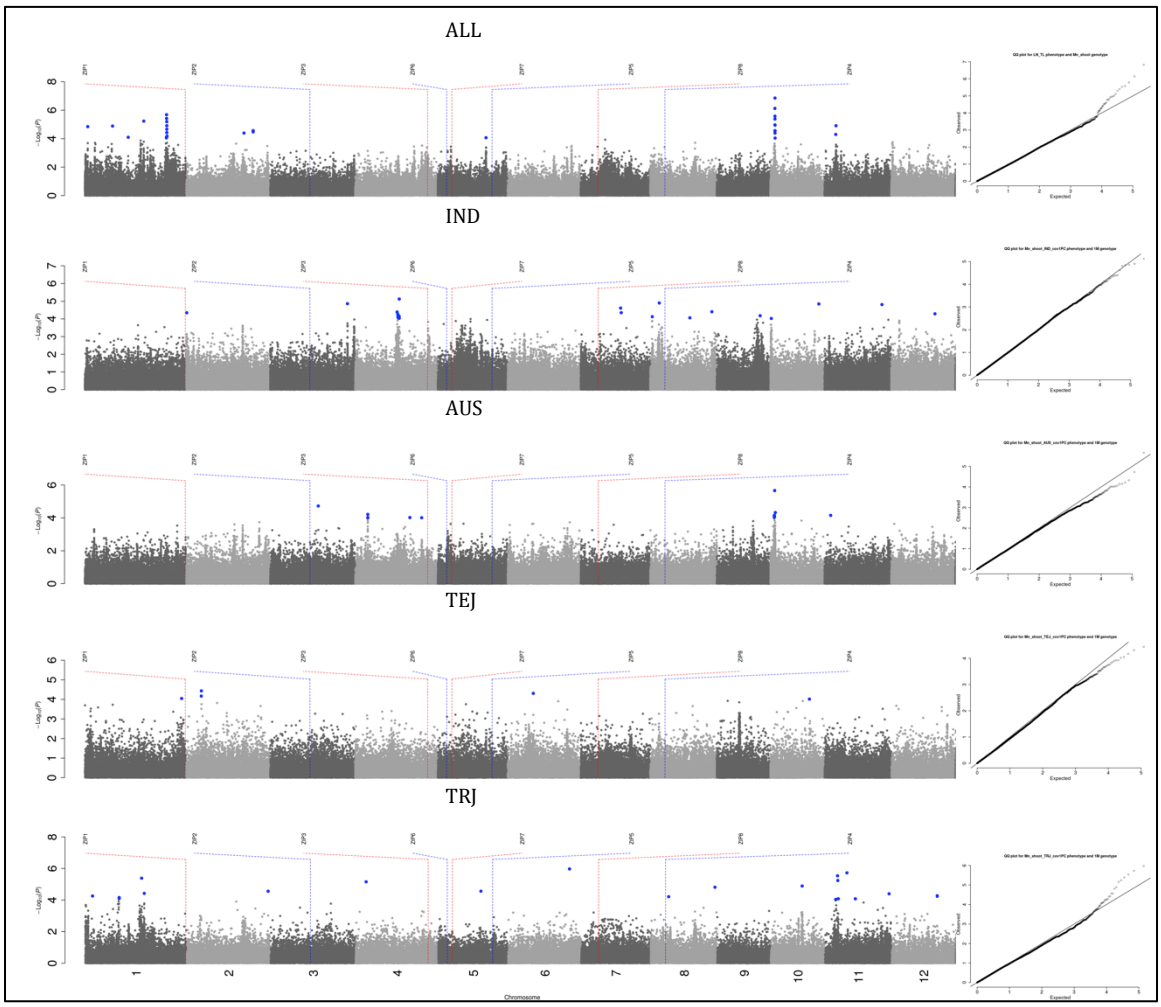
Supplementary Figure 3.25 Manhattan and quantile-quantile plots for magnesium (Mg) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



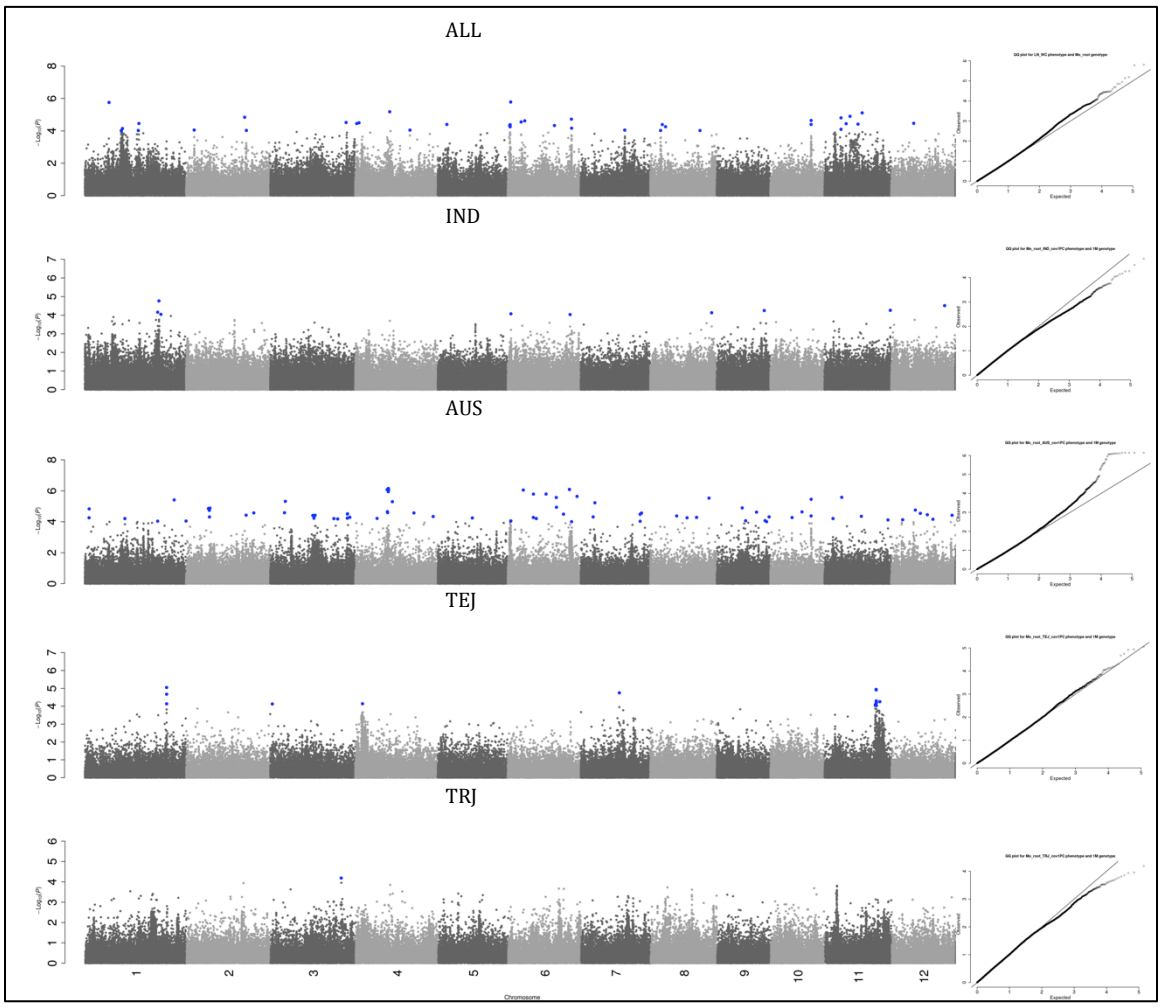
Supplementary Figure 3.26 Manhattan and quantile-quantile plots for magnesium (Mg) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



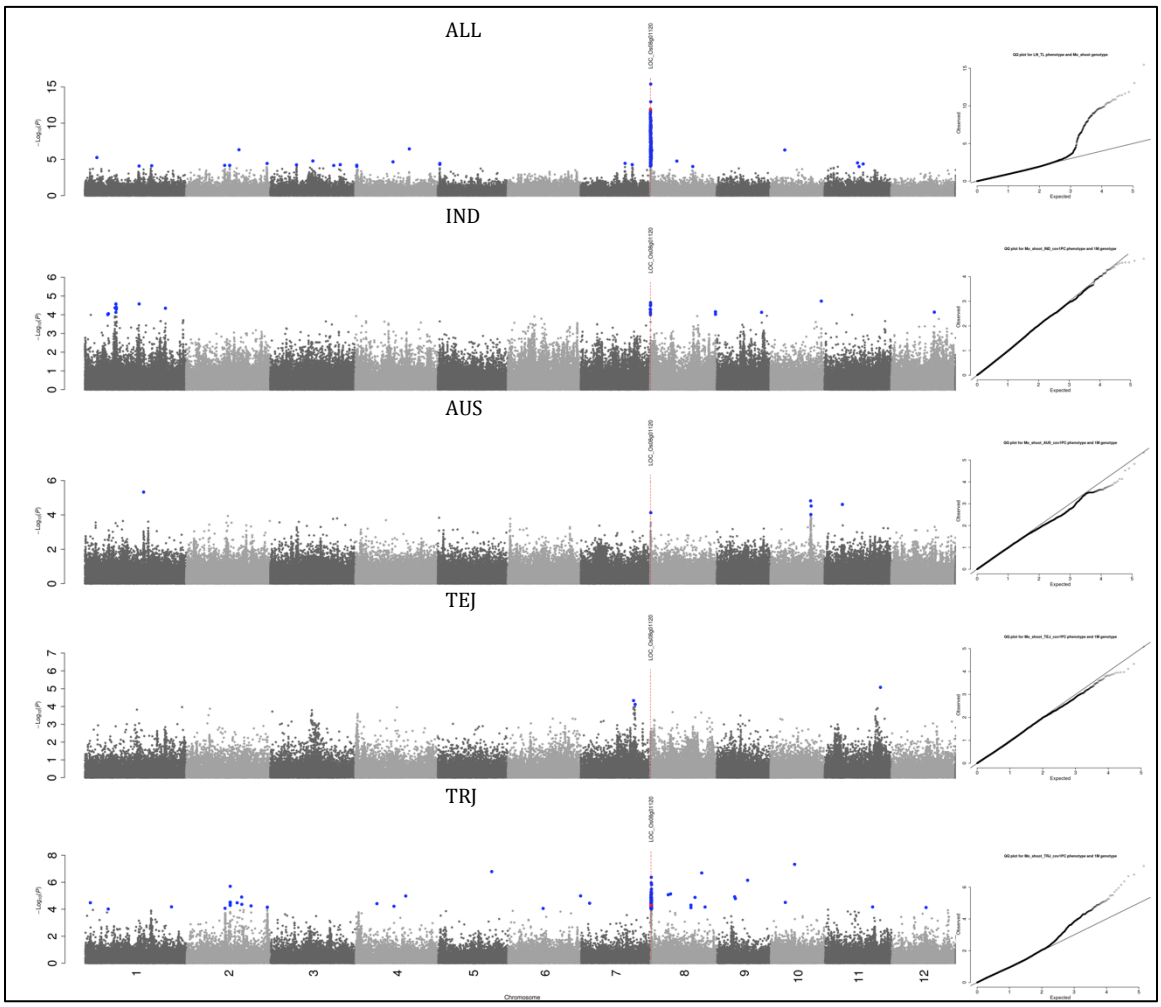
Supplementary Figure 3.27 Manhattan and quantile-quantile plots for manganese (Mn) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



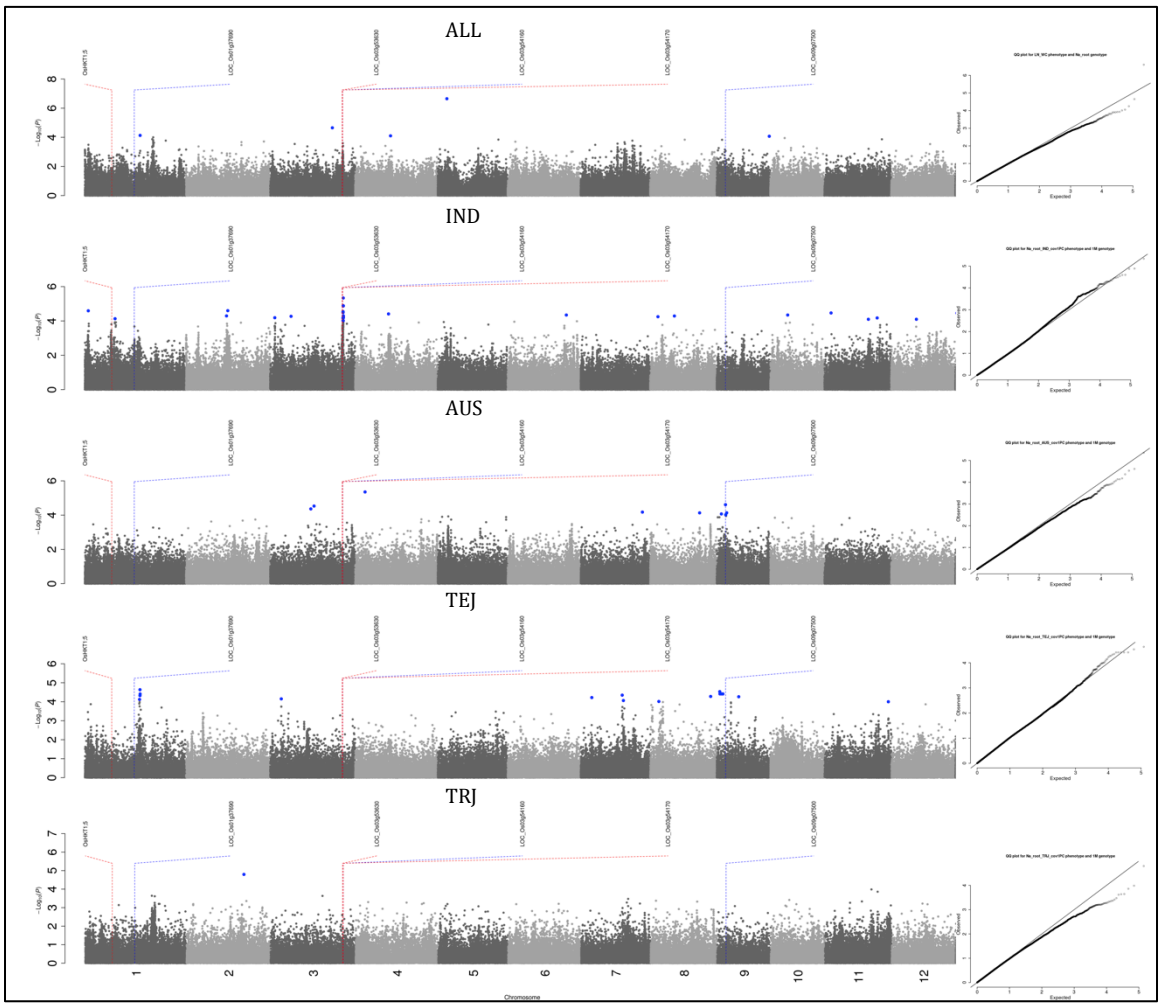
Supplementary Figure 3.28 Manhattan and quantile-quantile plots for manganese (Mn) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



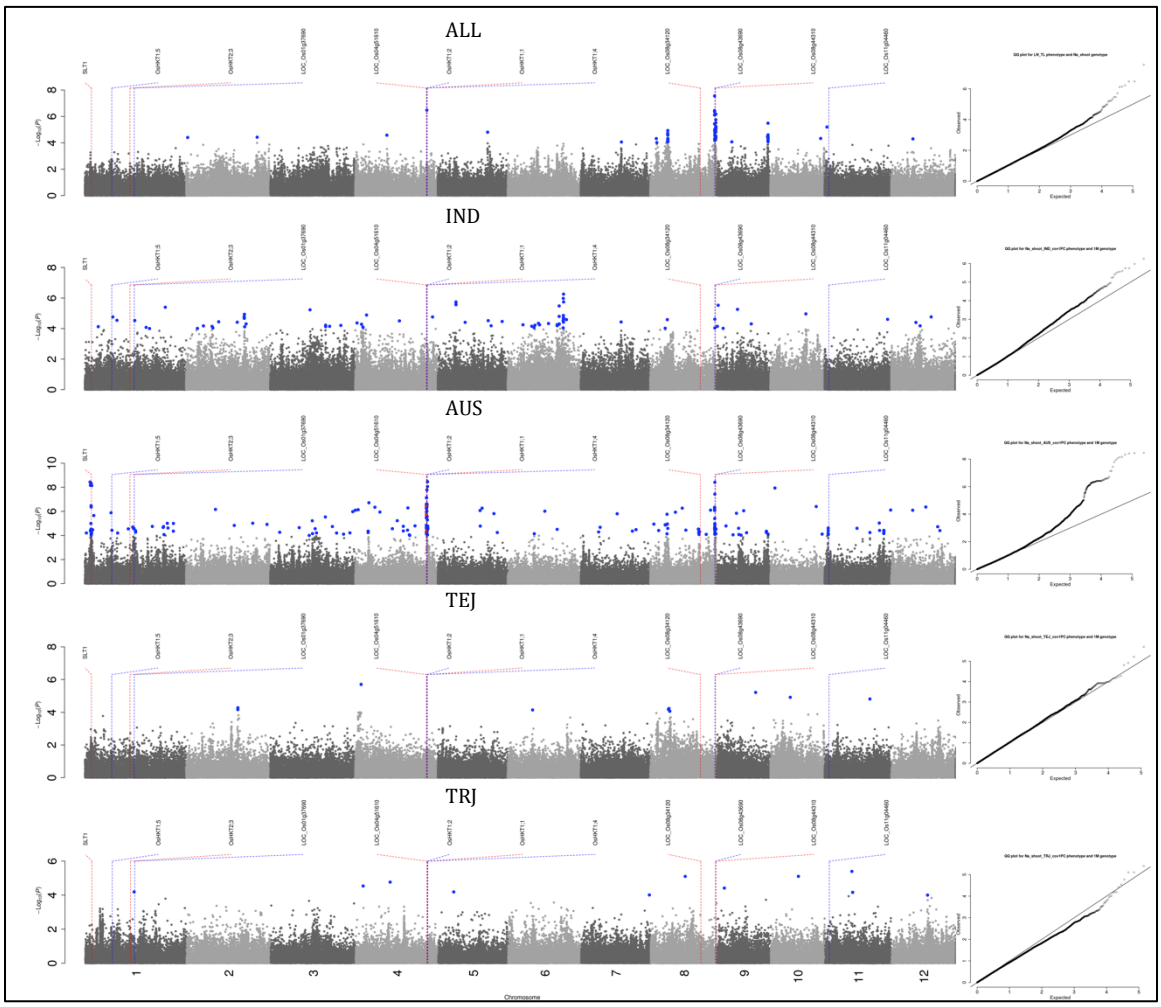
Supplementary Figure 3.29 Manhattan and quantile-quantile plots for molybdenum (Mo) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



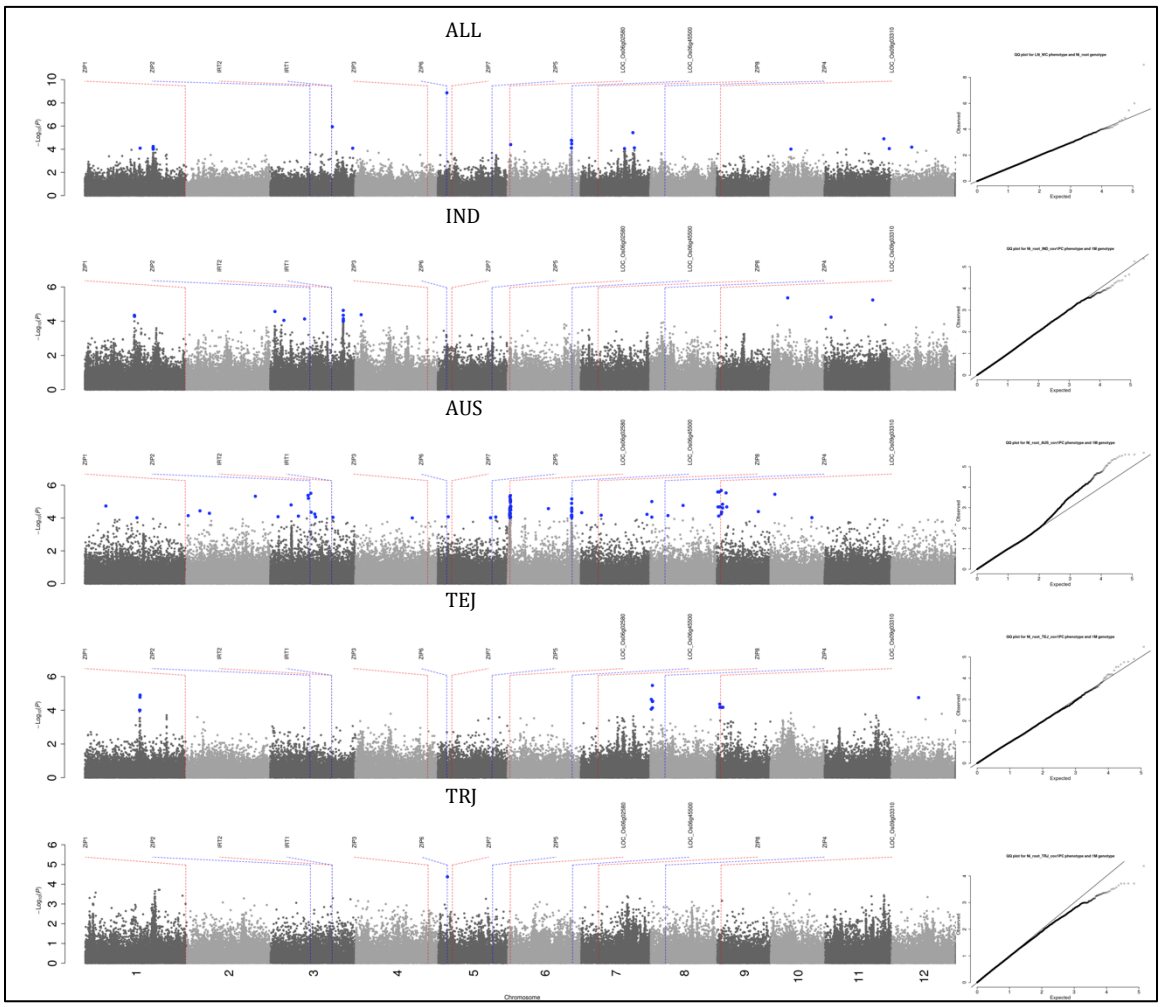
Supplementary Figure 3.30 Manhattan and quantile-quantile plots for molybdenum (Mo) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



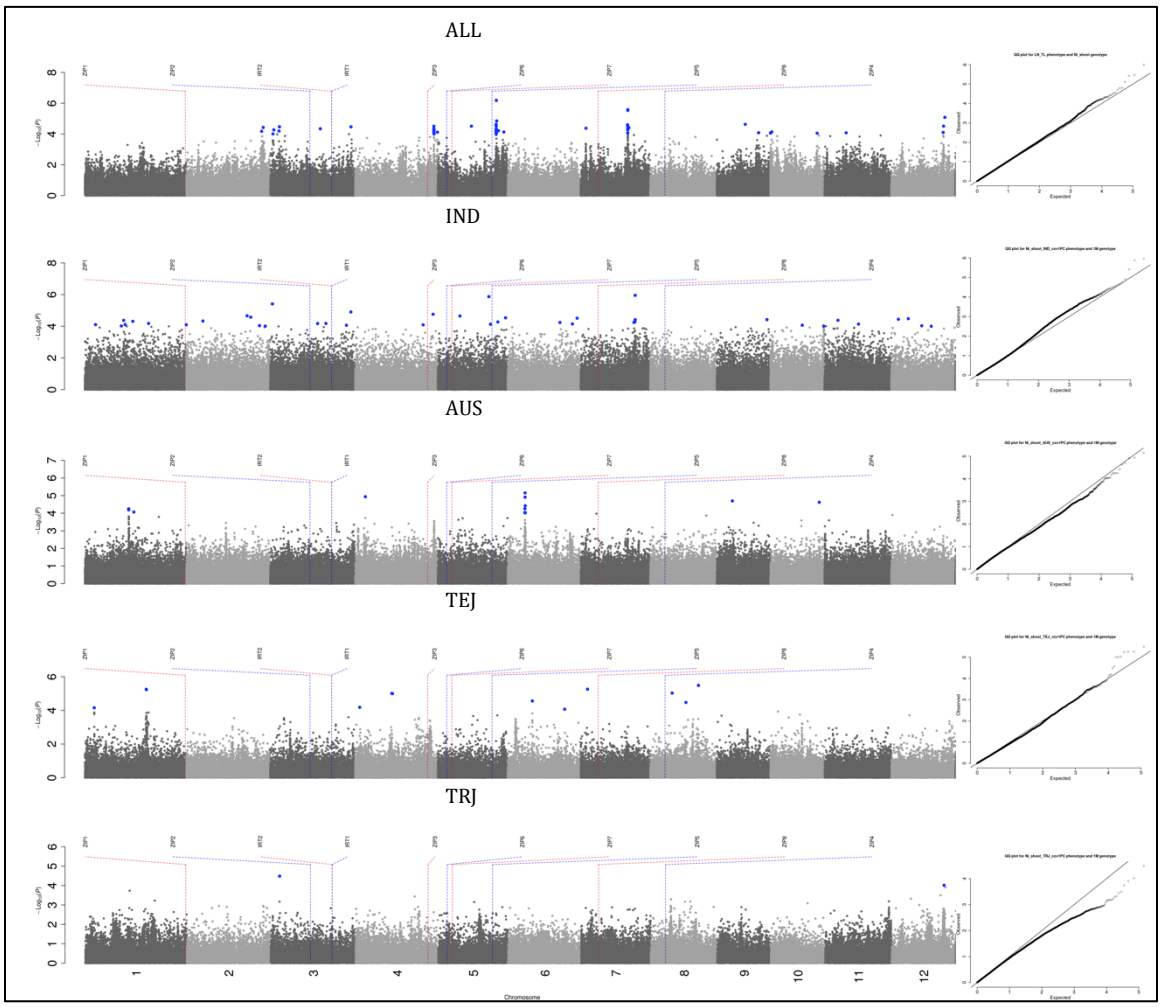
Supplementary Figure 3.31 Manhattan and quantile-quantile plots for sodium (Na) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



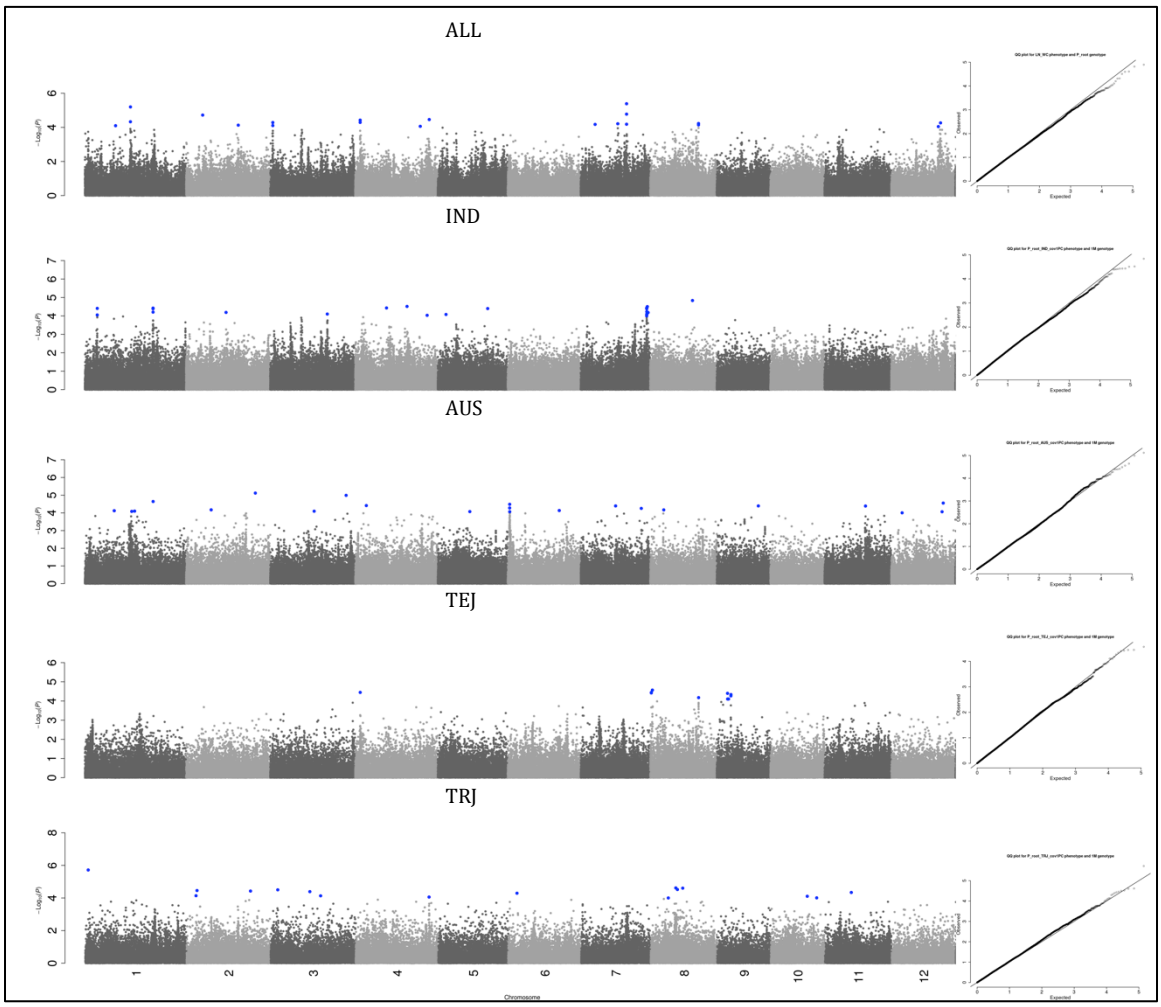
Supplementary Figure 3.32 Manhattan and quantile-quantile plots for sodium (Na) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



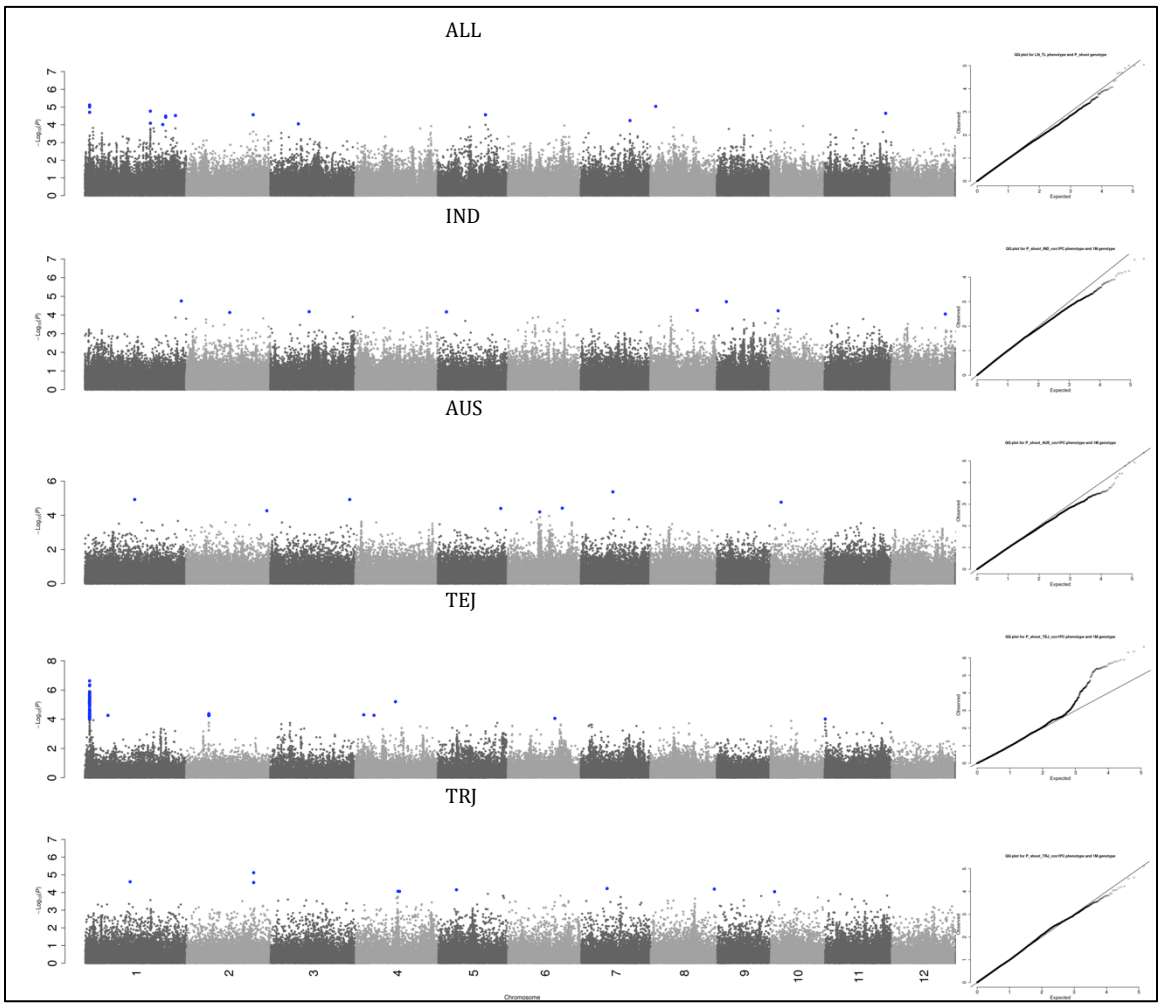
Supplementary Figure 3.33 Manhattan and quantile-quantile plots for nickel (Ni) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



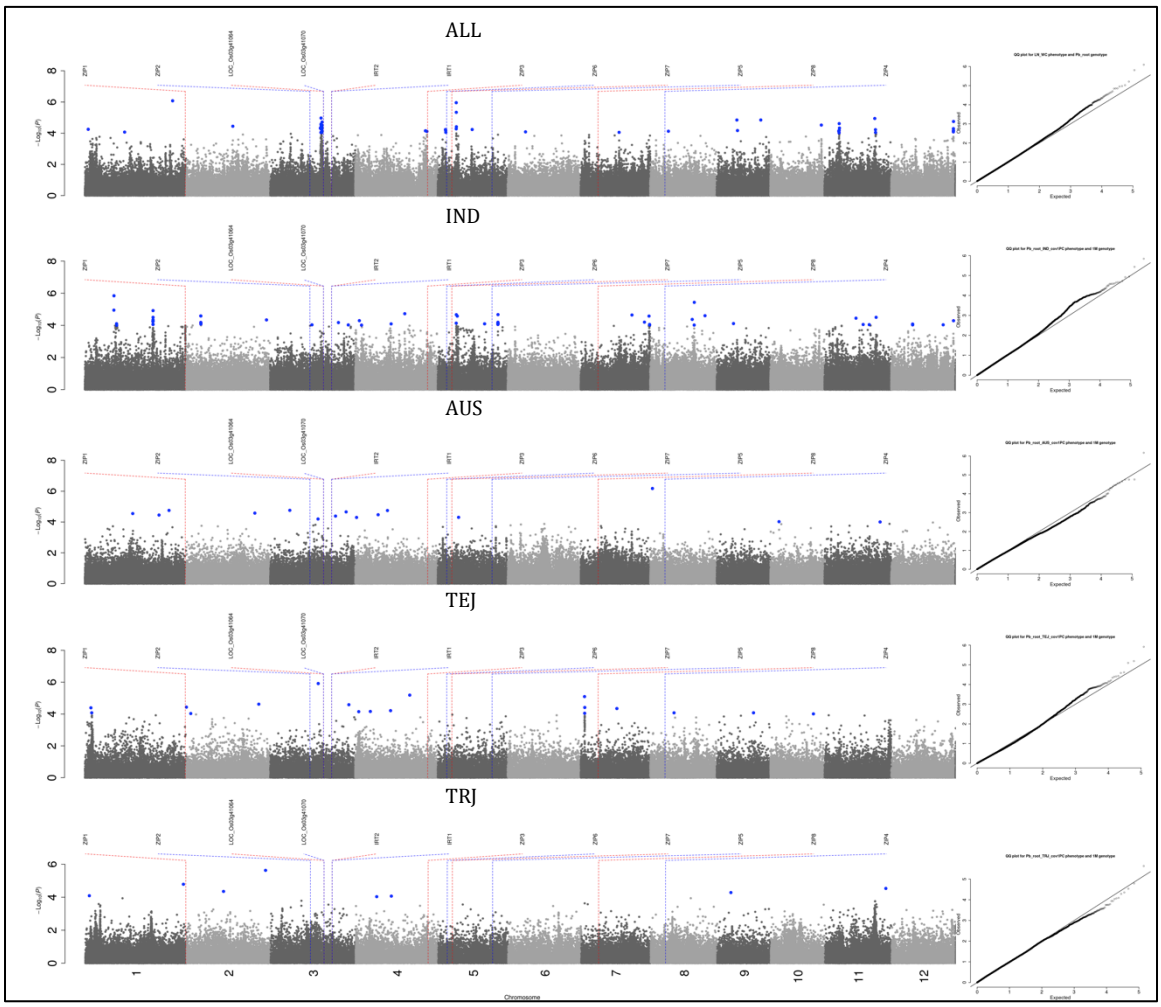
Supplementary Figure 3.34 Manhattan and quantile-quantile plots for nickel (Ni) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



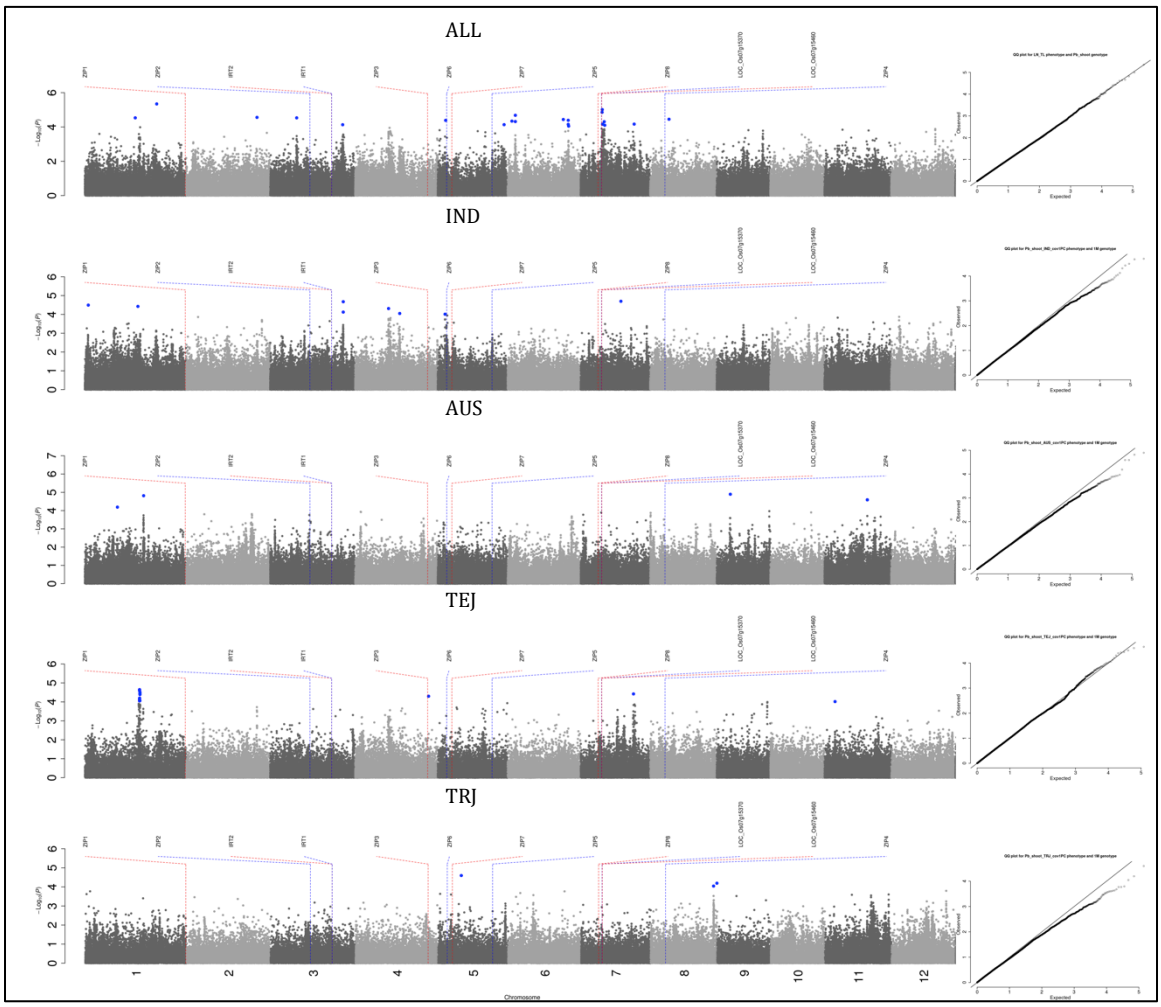
Supplementary Figure 3.35 Manhattan and quantile-quantile plots for phosphorus (P) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



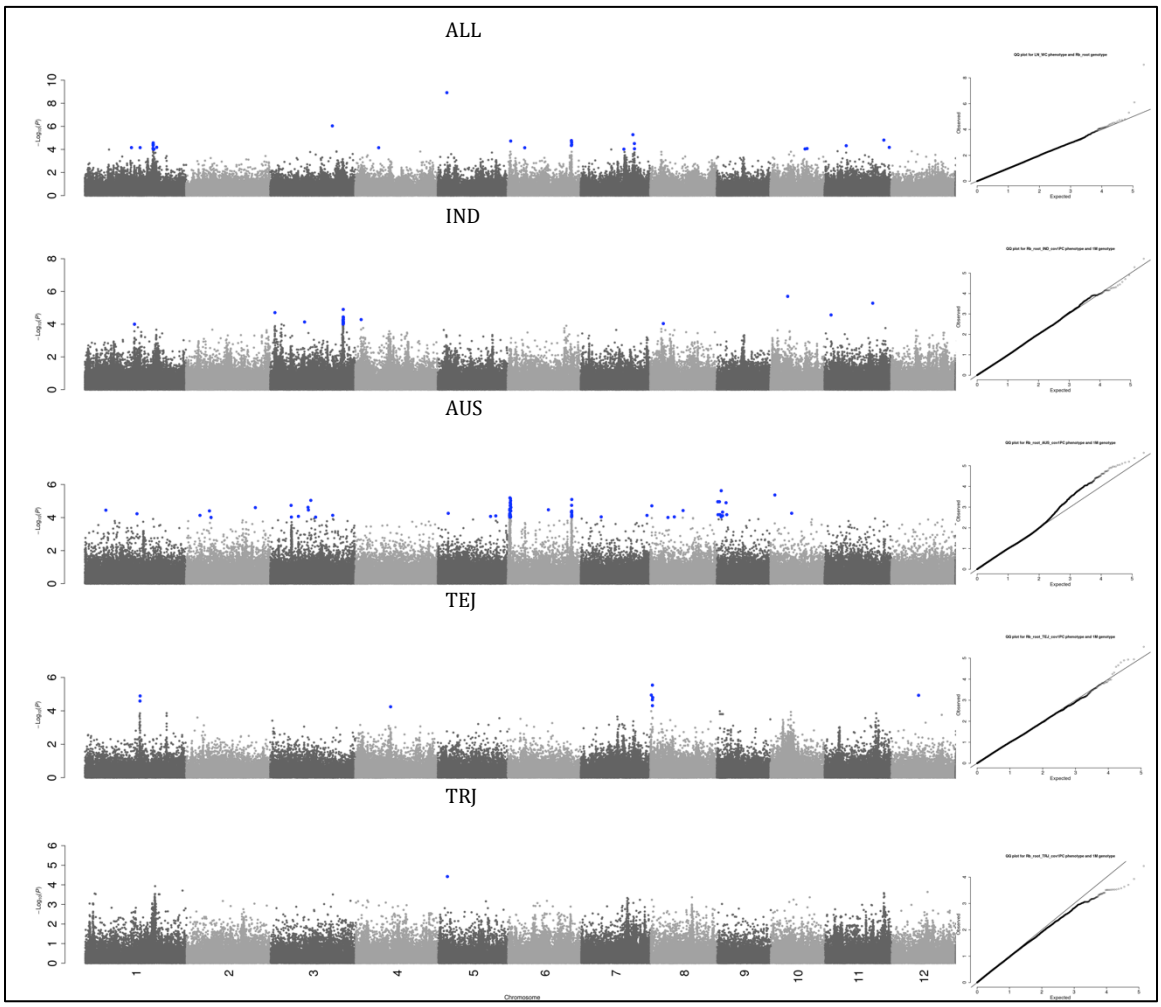
Supplementary Figure 3.36 Manhattan and quantile-quantile plots for phosphorus (P) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



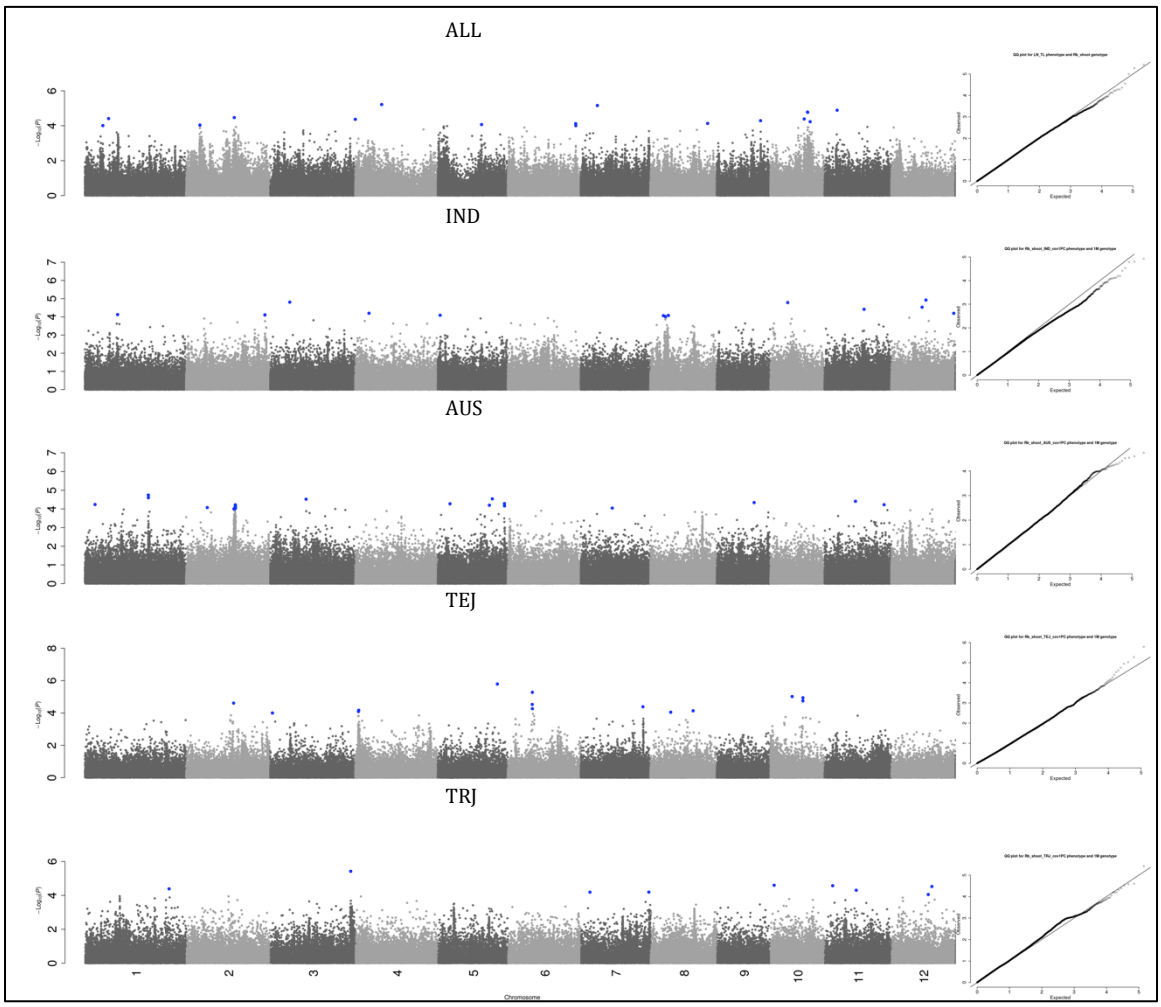
Supplementary Figure 3.37 Manhattan and quantile-quantile plots for lead (Pb) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



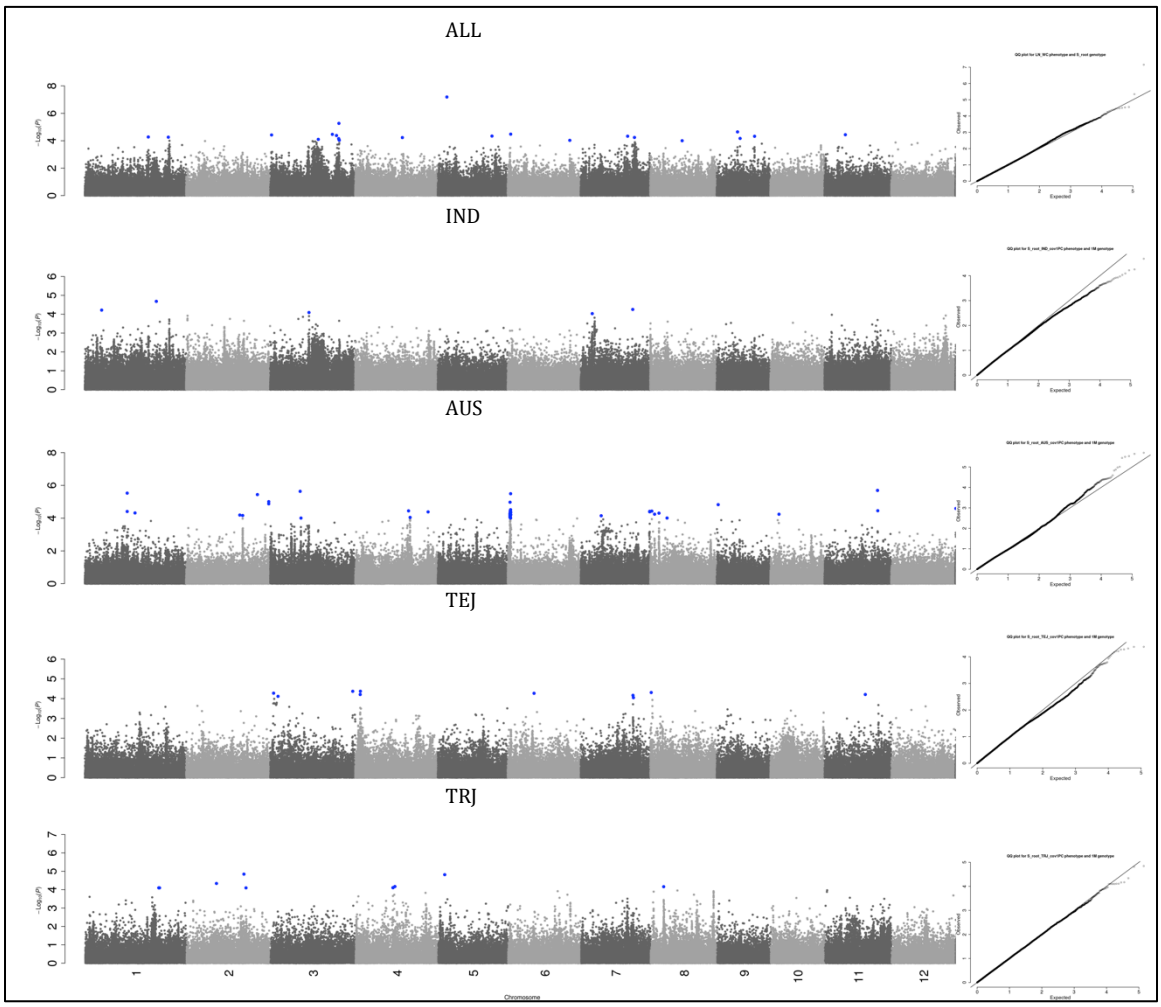
Supplementary Figure 3.38 Manhattan and quantile-quantile plots for lead (Pb) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



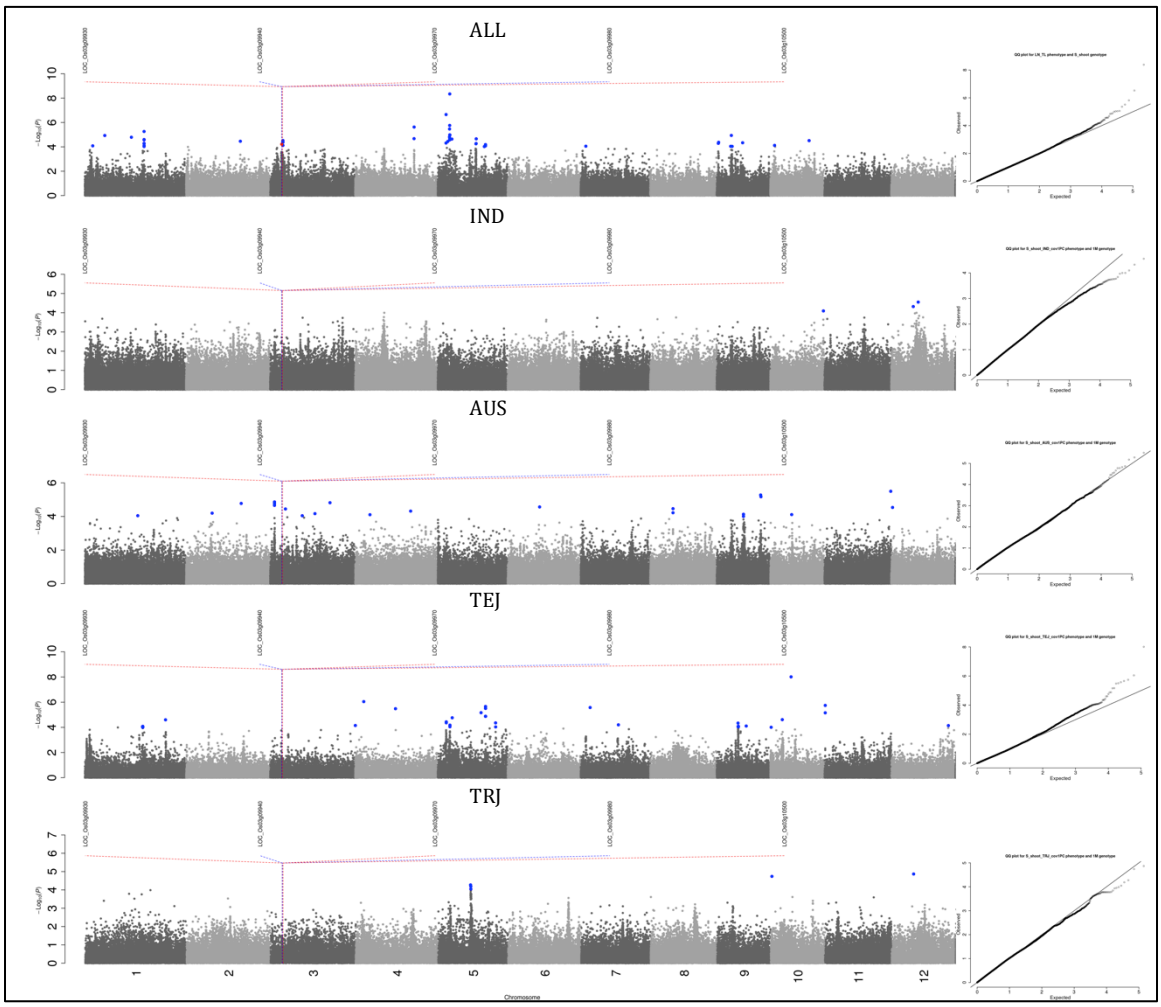
Supplementary Figure 3.39 Manhattan and quantile-quantile plots for rubidium (Rb) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



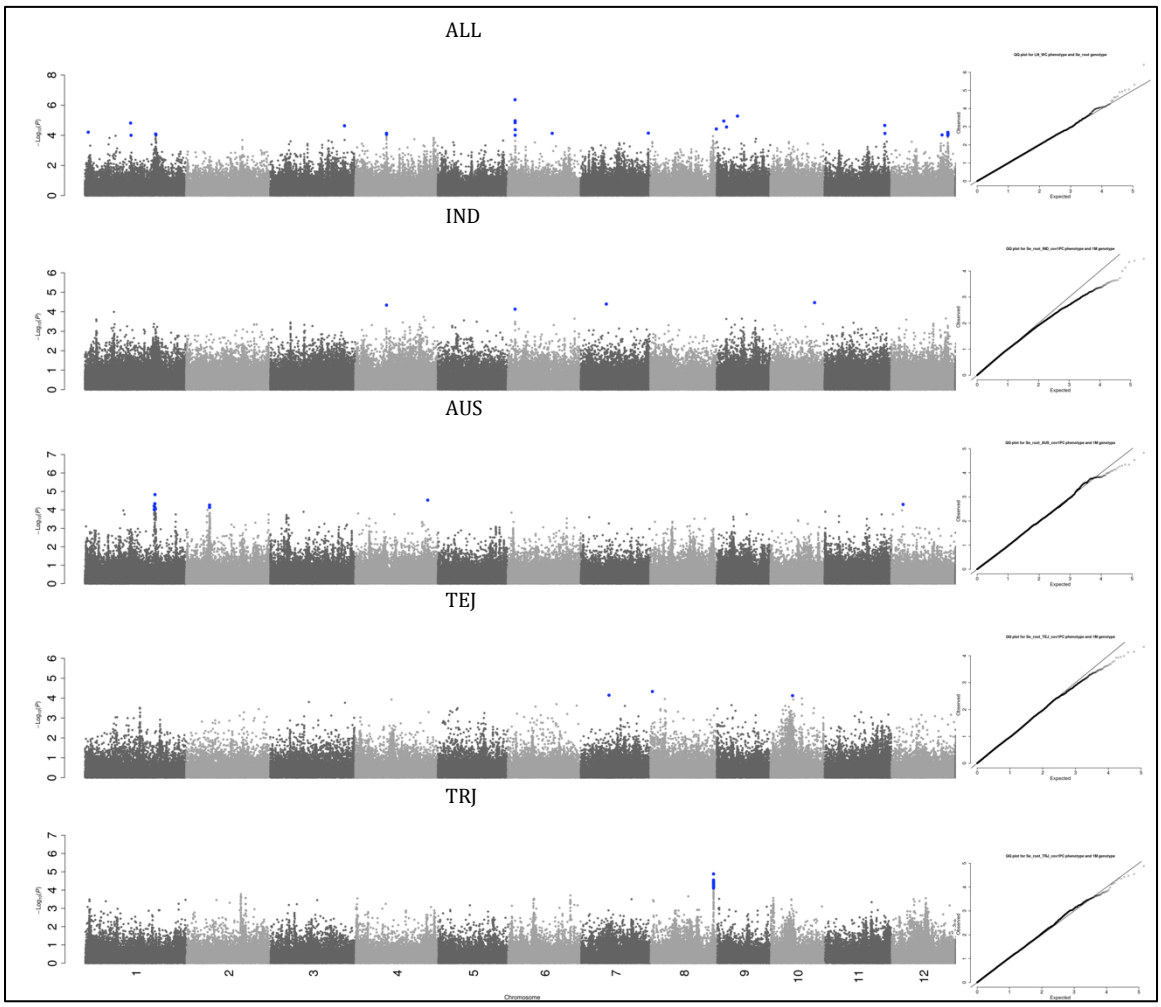
Supplementary Figure 3.40 Manhattan and quantile-quantile plots for rubidium (Rb) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



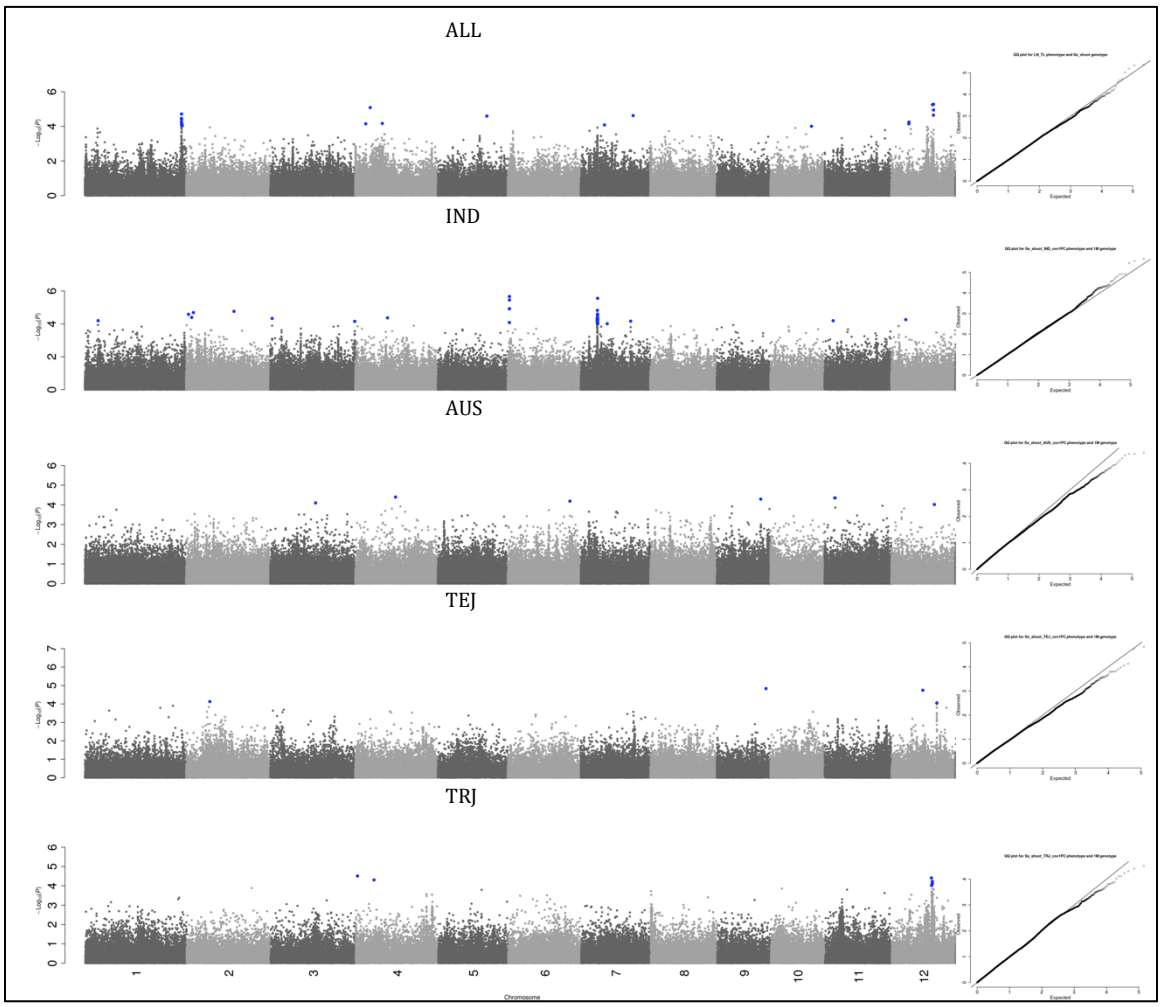
Supplementary Figure 3.41 Manhattan and quantile-quantile plots for sulfur (S) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



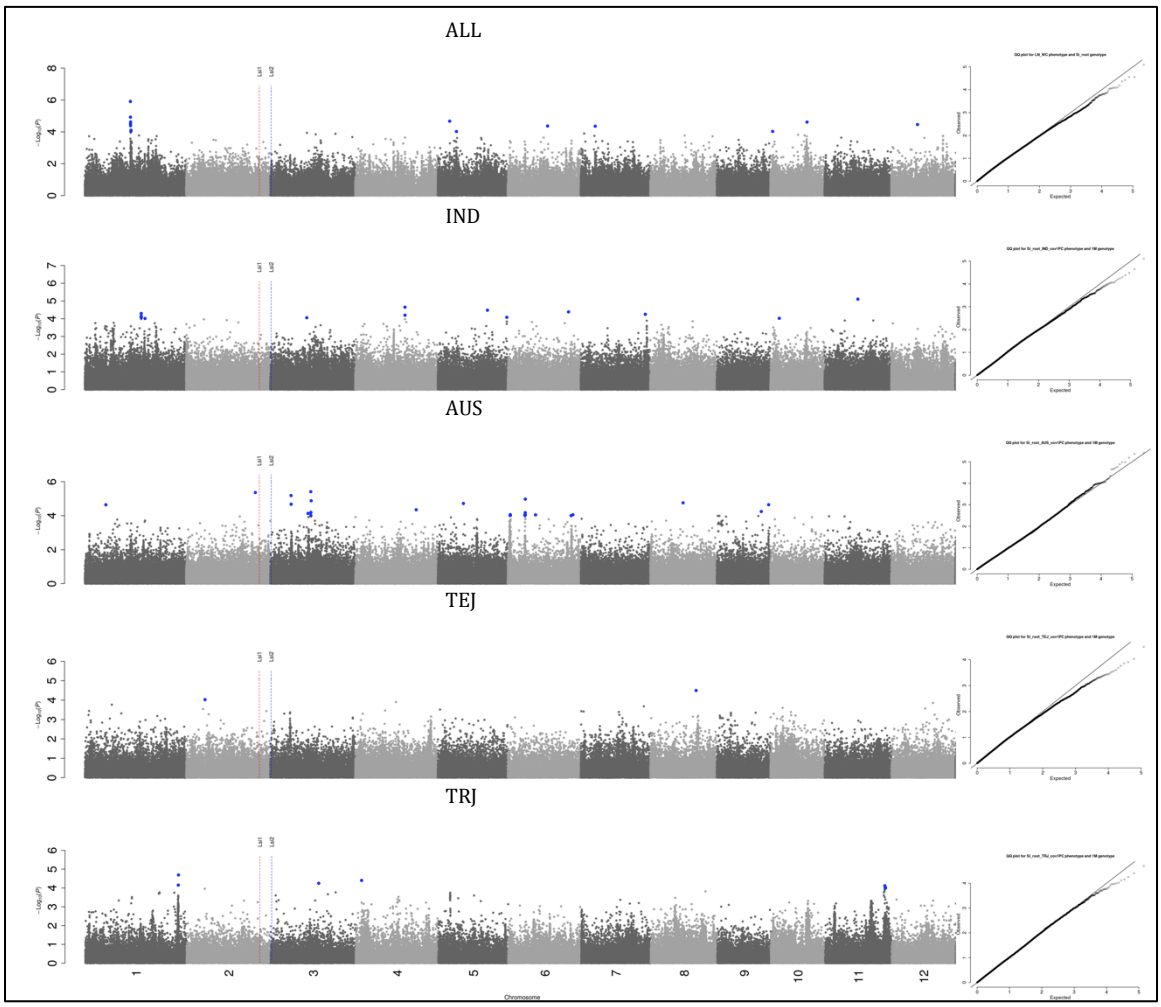
Supplementary Figure 3.42 Manhattan and quantile-quantile plots for sulfur (S) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



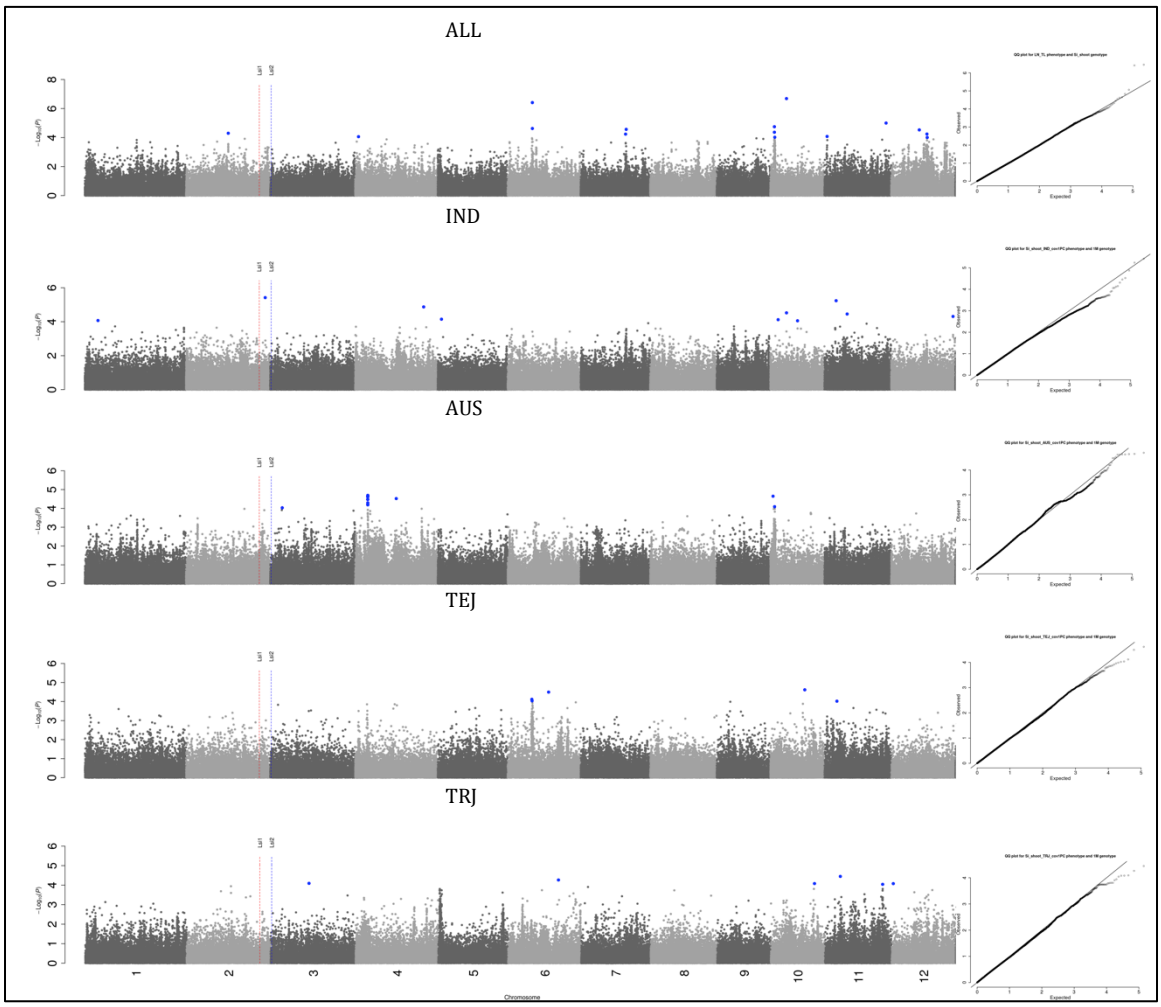
Supplementary Figure 3.43 Manhattan and quantile-quantile plots for selenium (Se) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



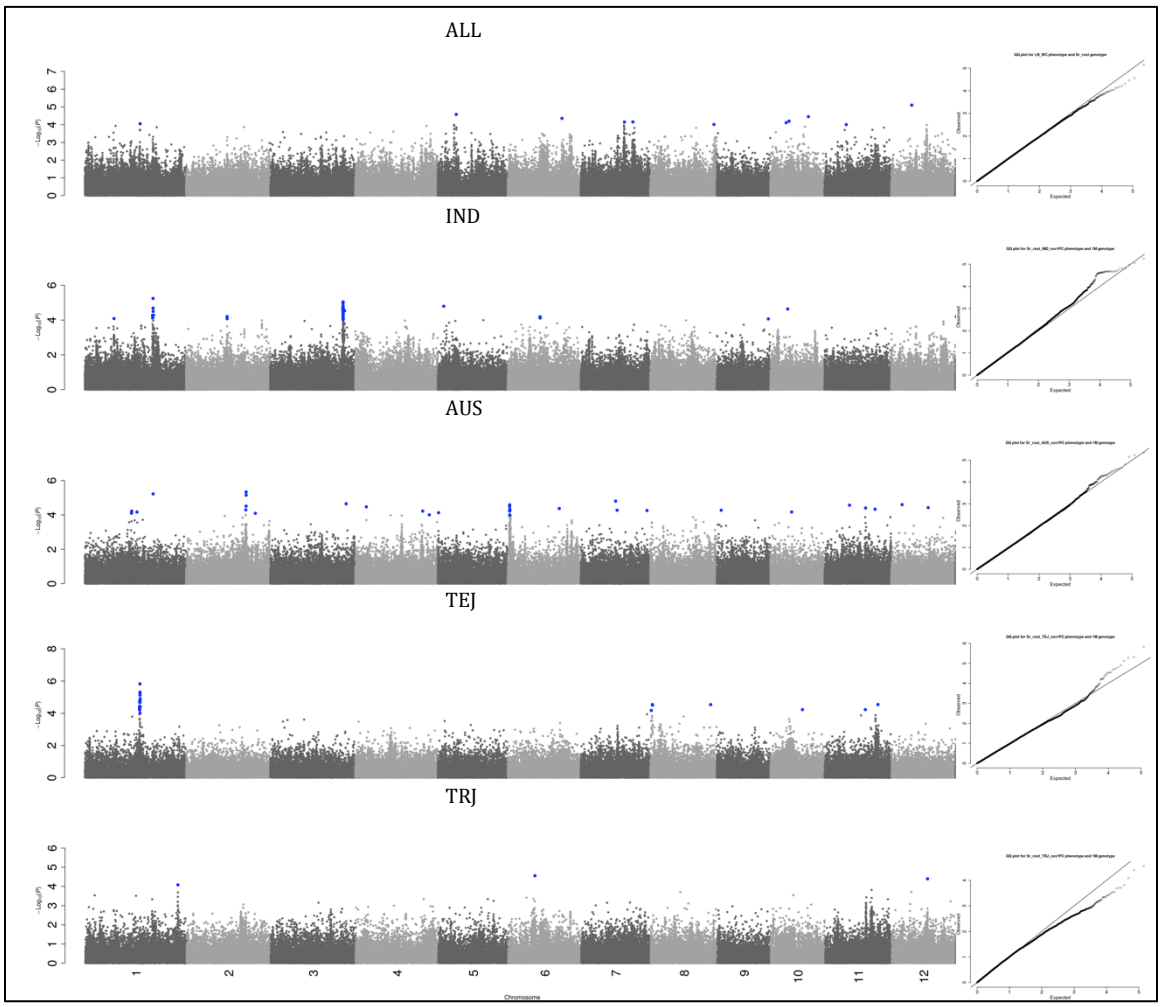
Supplementary Figure 3.44 Manhattan and quantile-quantile plots for selenium (Se) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



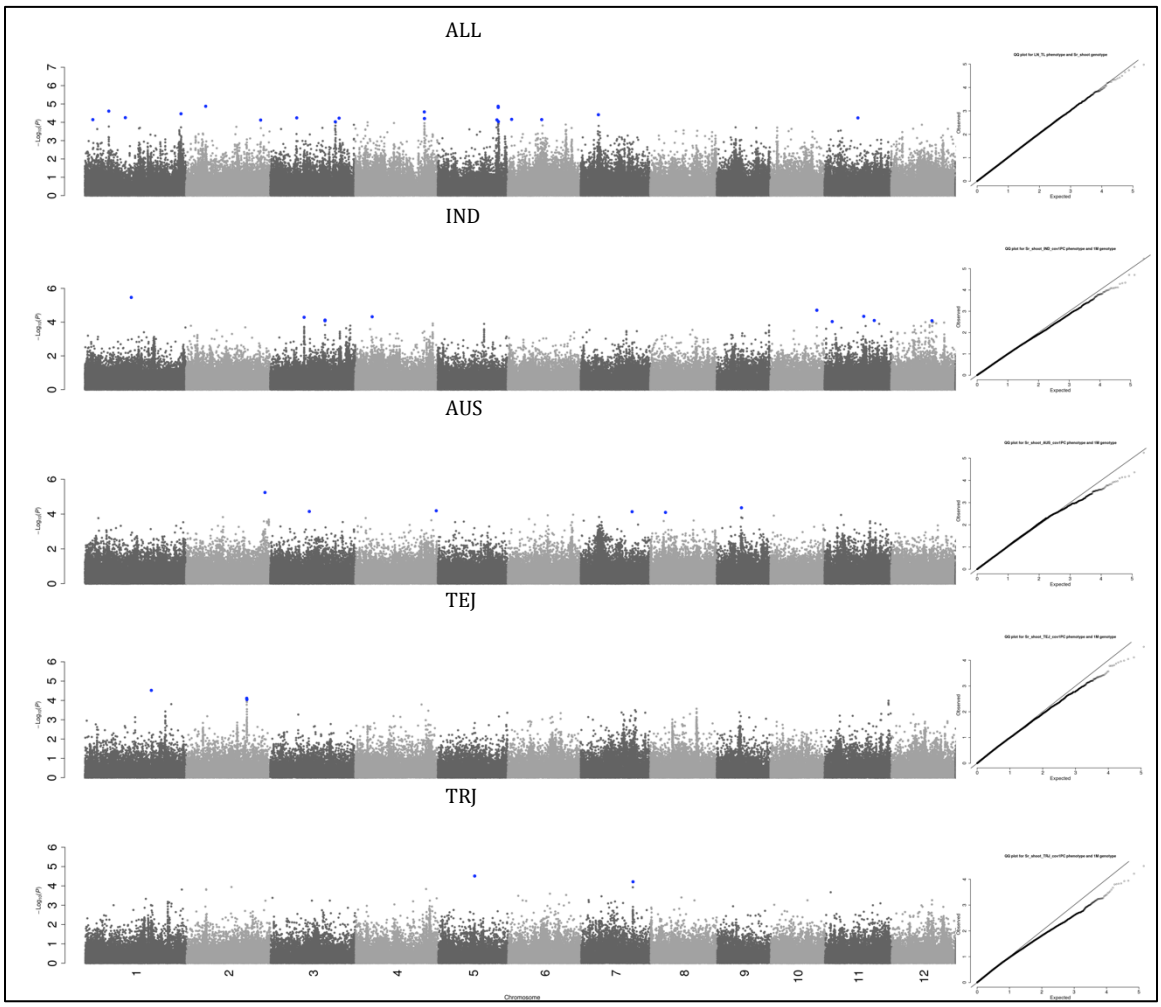
Supplementary Figure 3.45 Manhattan and quantile-quantile plots for silicon (Si) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



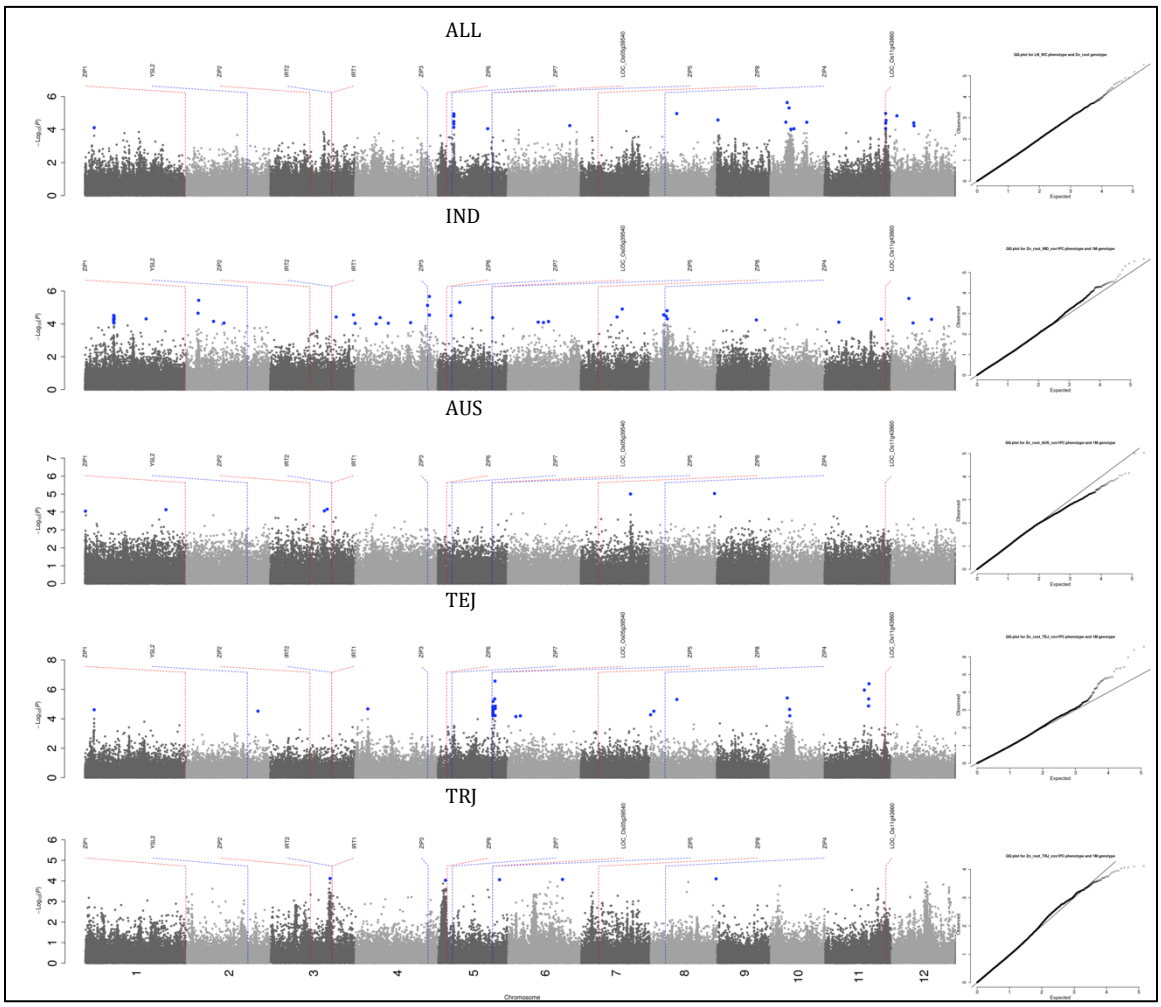
Supplementary Figure 3.46 Manhattan and quantile-quantile plots for silicon (Si) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



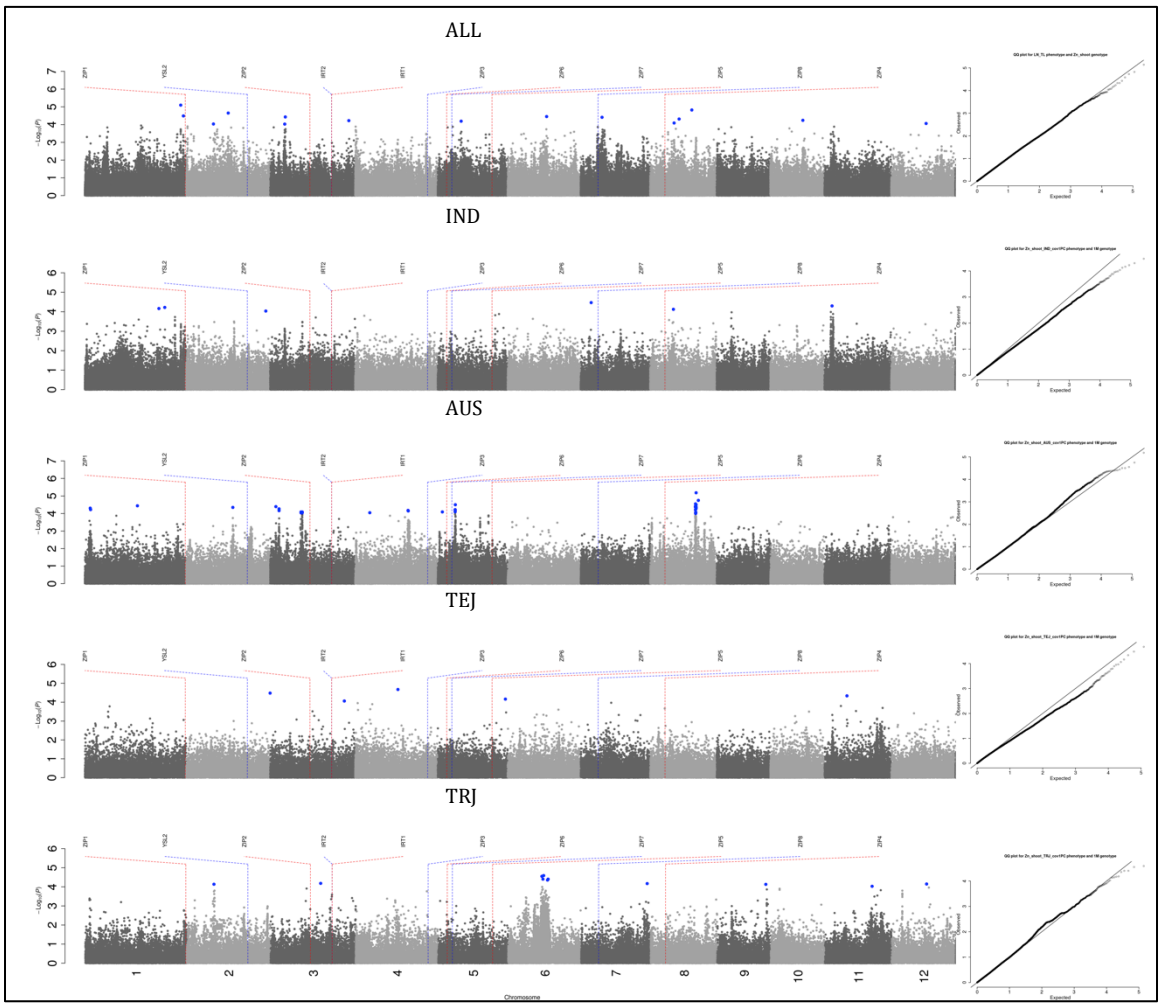
Supplementary Figure 3.47 Manhattan and quantile-quantile plots for strontium (Sr) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



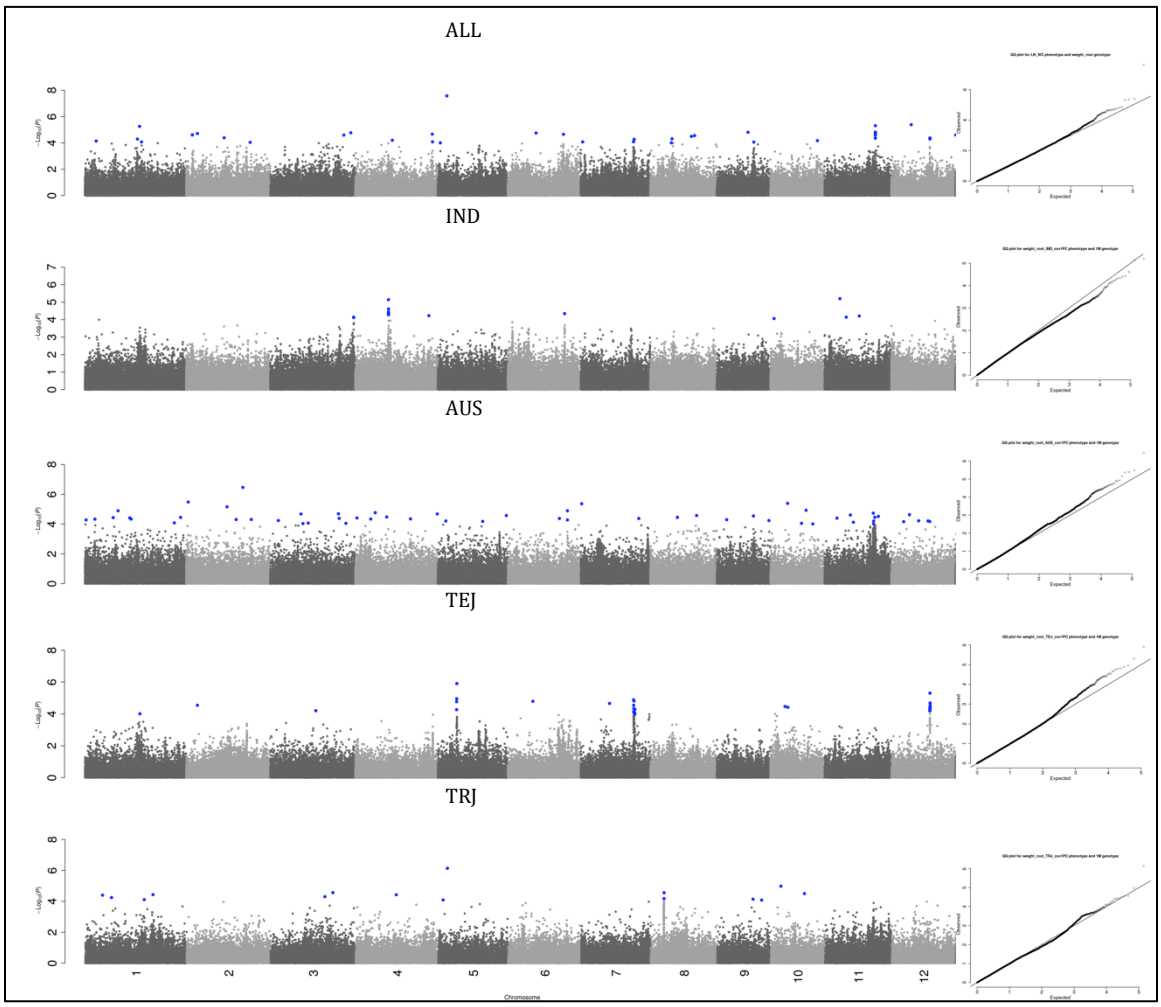
Supplementary Figure 3.48 Manhattan and quantile-quantile plots for strontium (Sr) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



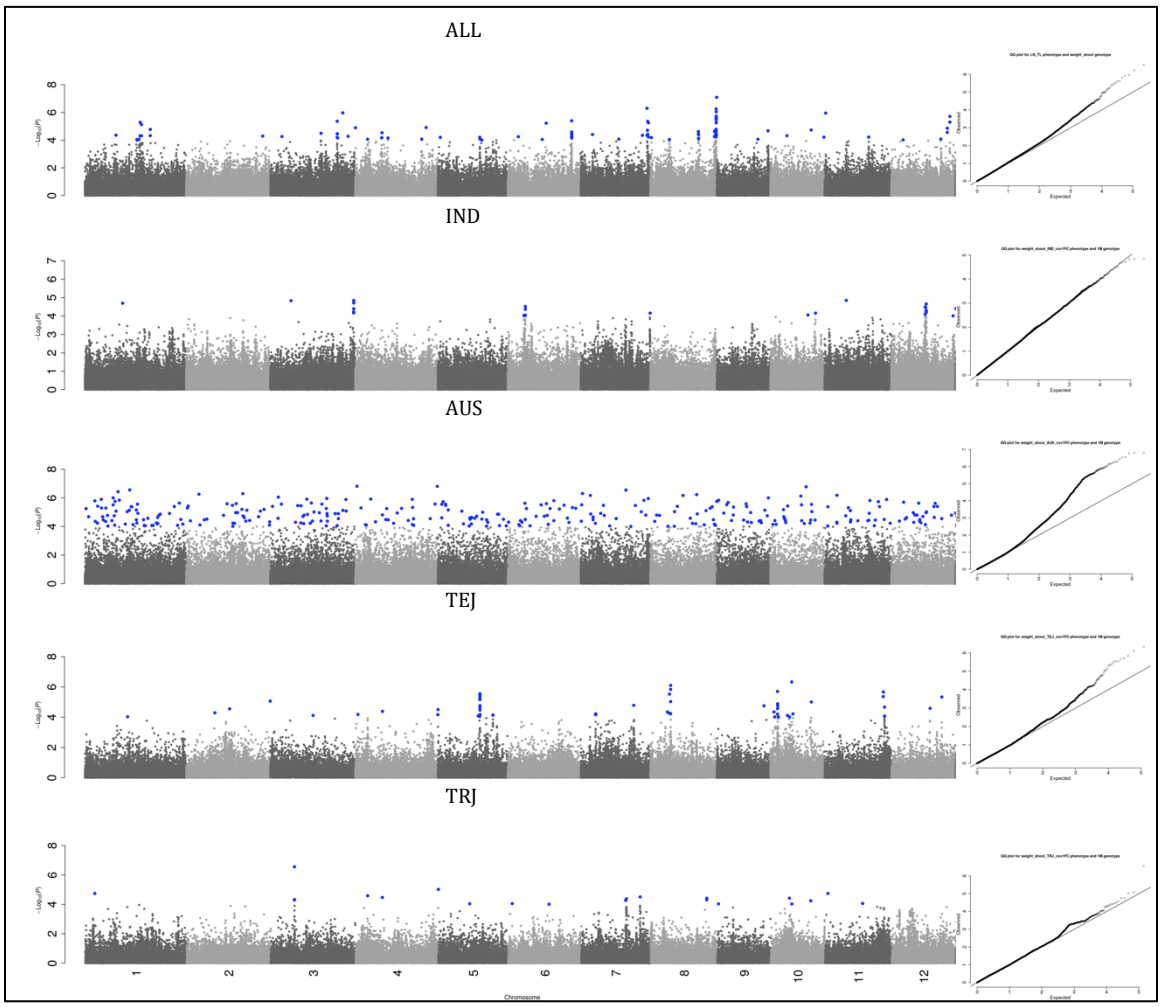
Supplementary Figure 3.49 Manhattan and quantile-quantile plots for zinc (Zn) root content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



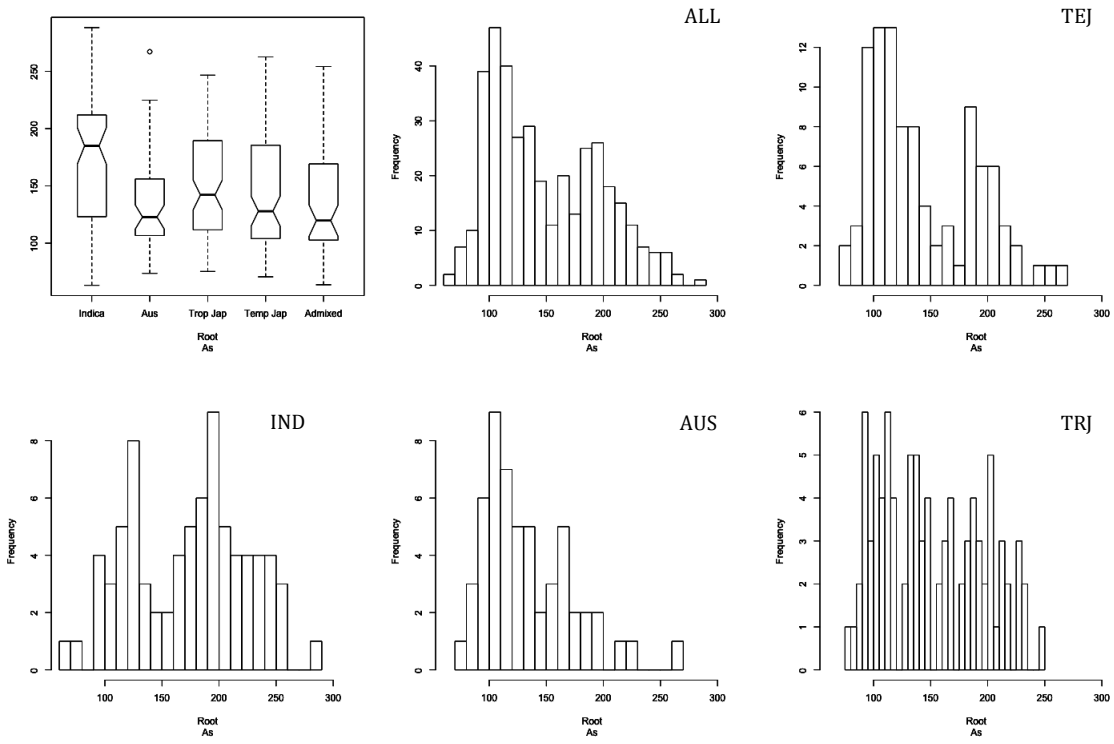
Supplementary Figure 3.50 Manhattan and quantile-quantile plots for zinc (Zn) shoot content. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



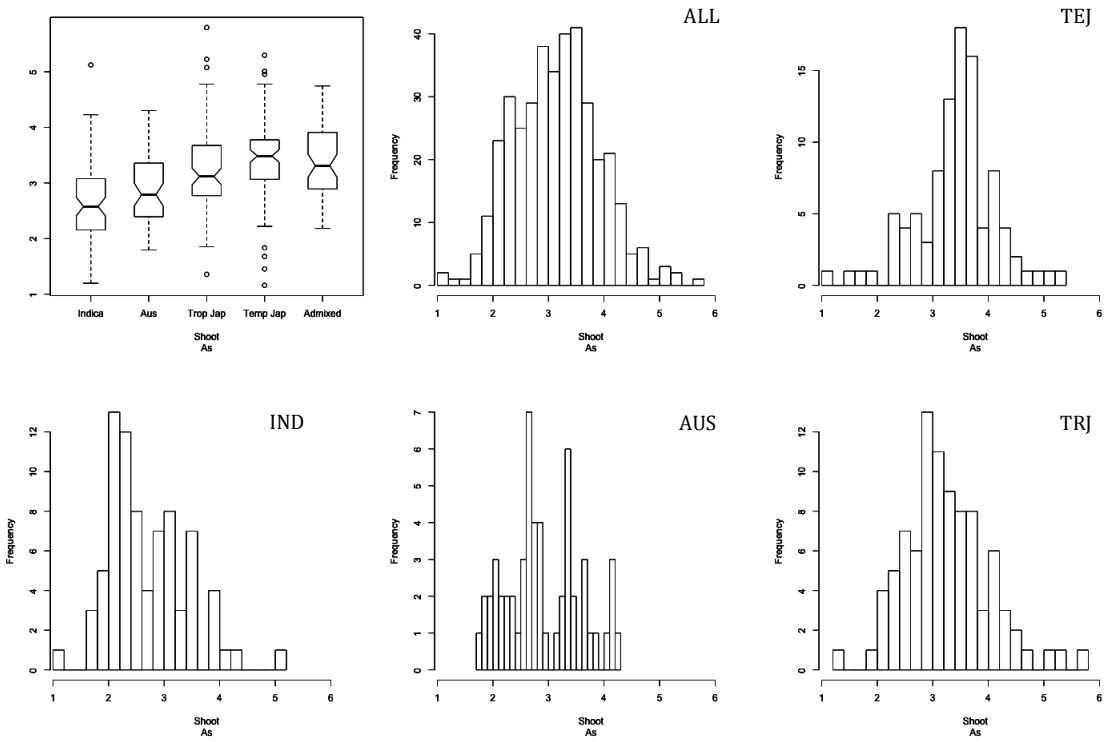
Supplementary Figure 3.51 Manhattan and quantile-quantile plots for whole root dry weight biomass. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



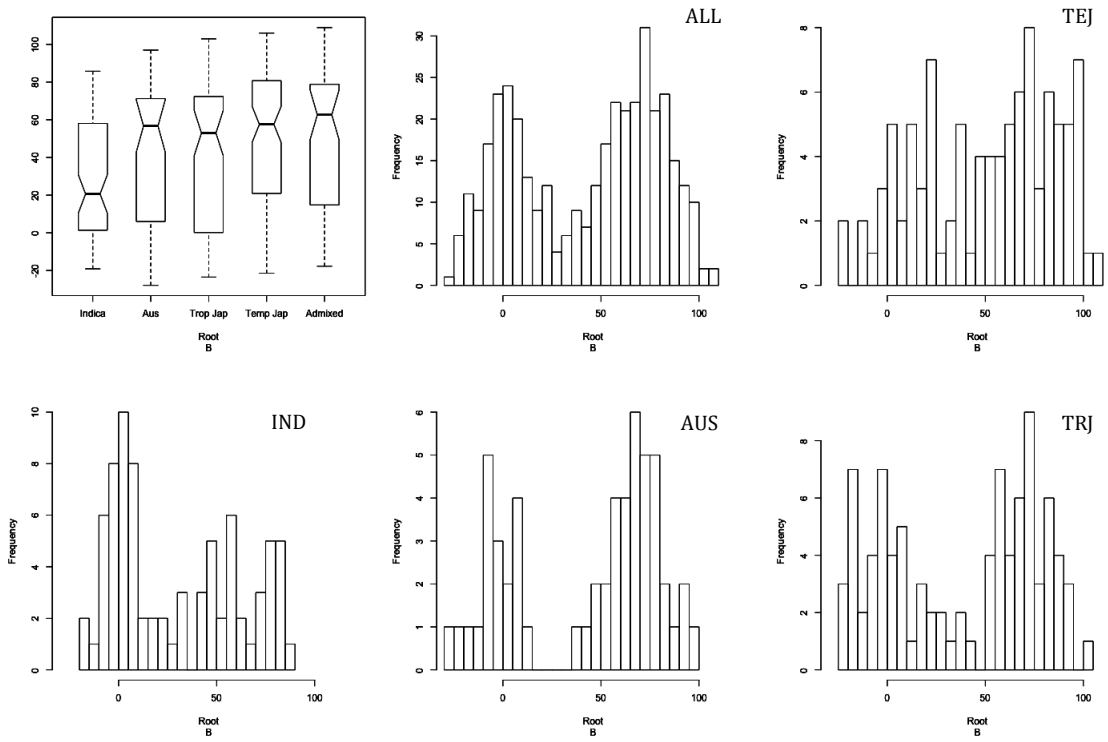
Supplementary Figure 3.52 Manhattan and quantile-quantile plots for whole shoot dry weight biomass. Each subpopulation specific analysis is displayed as well as ALL. Blue SNPs have a p-value of less than 10^{-4}, and red SNPs (if any) are found within gene models for candidate genes. Only cloned candidate genes and putative candidate genes found within peaks are shown (if applicable).



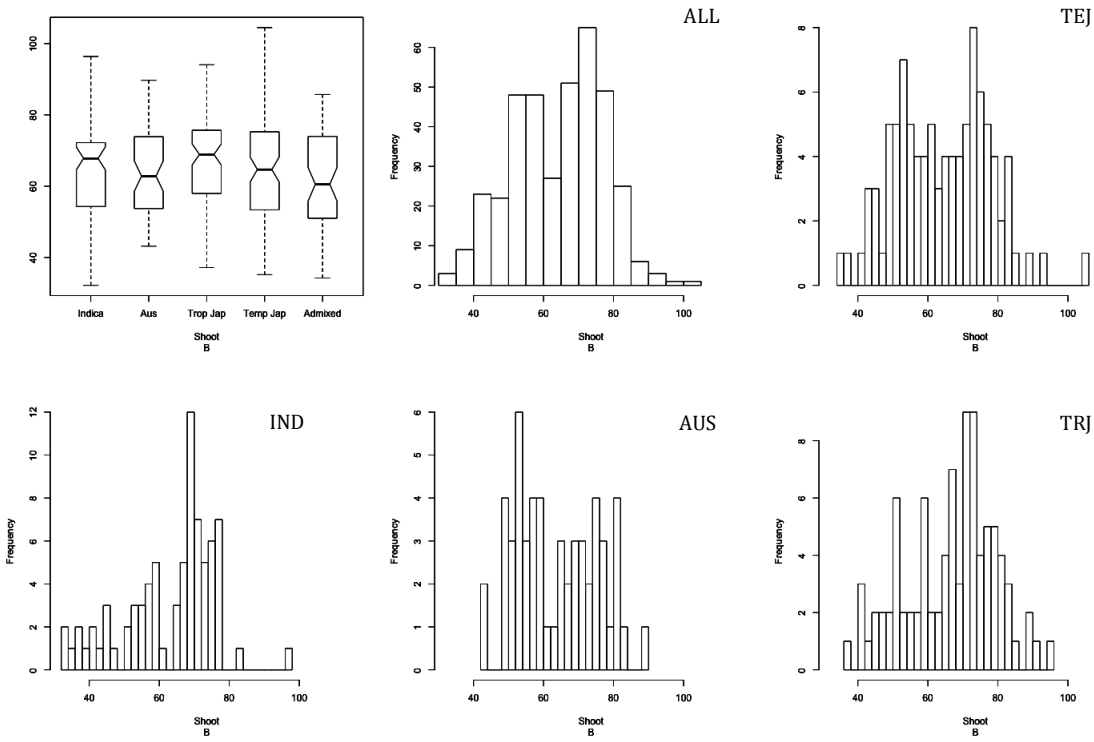
Supplementary Figure 3.53 Boxplot and histogram distributions of arsenic (As) root phenotype data by subpopulation. Subpopulation structure explained 4% of the phenotypic variance (p-value < 0.0001).



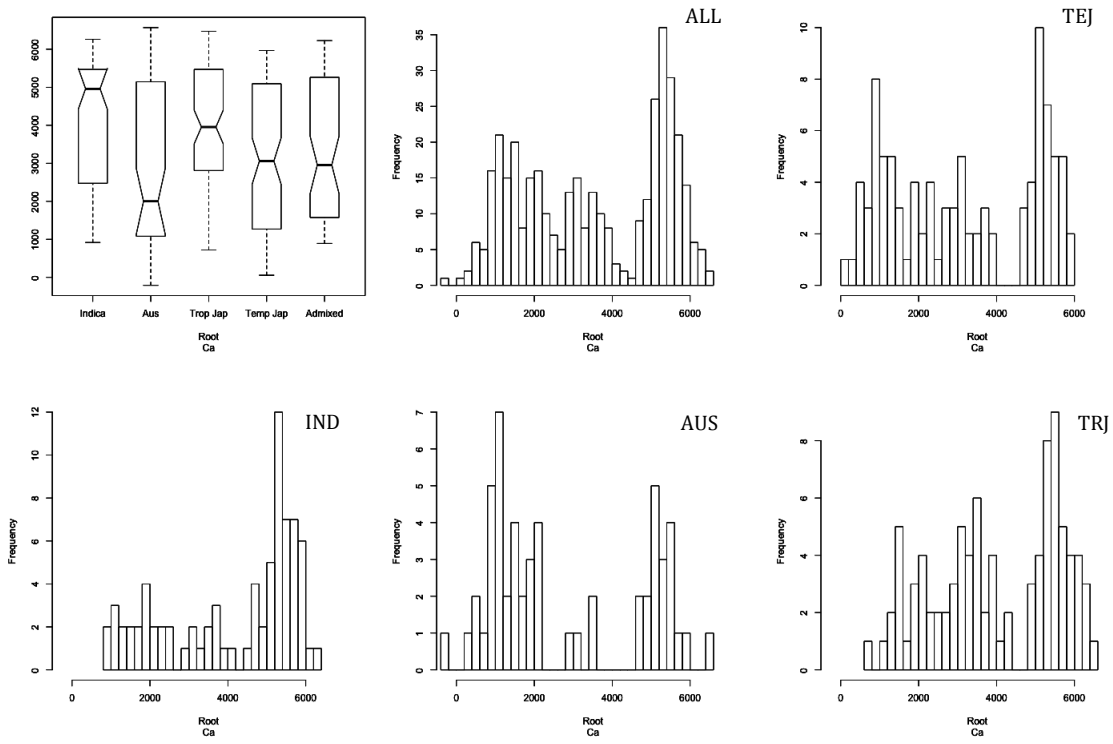
Supplementary Figure 3.54 Boxplot and histogram distributions of arsenic (As) shoot phenotype data by subpopulation. Subpopulation structure explained 12% of the phenotypic variance (p-value < 0.0001).



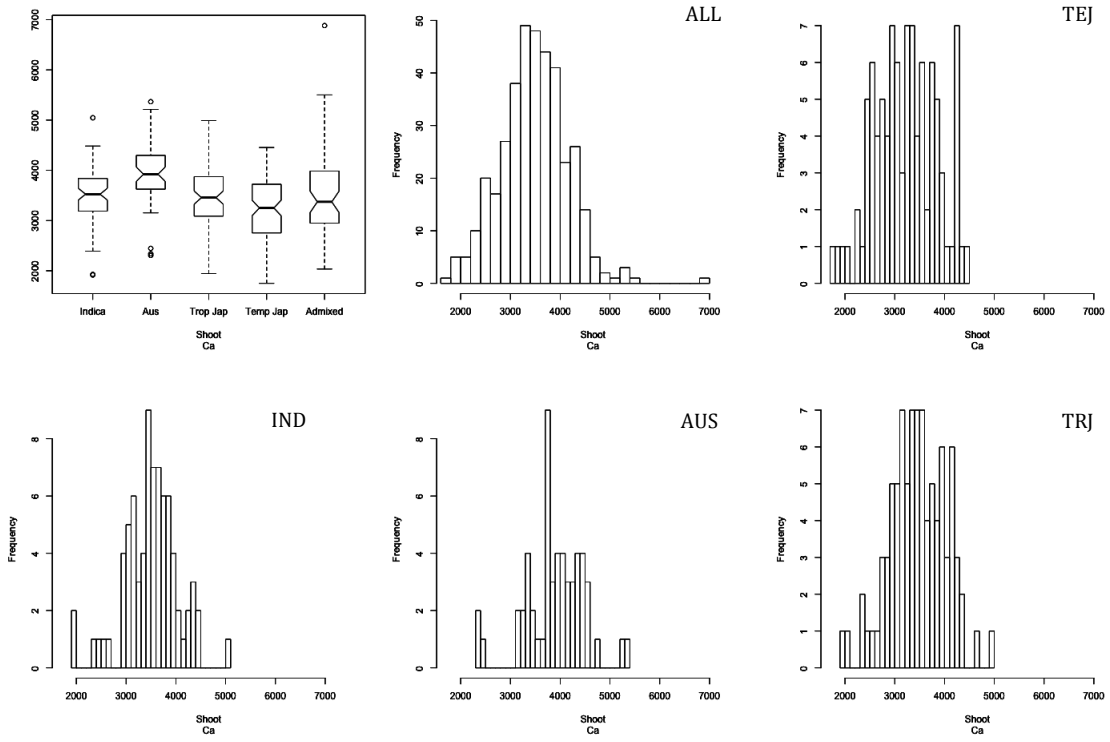
Supplementary Figure 3.55 Boxplot and histogram distributions of boron (B) root phenotype data by subpopulation. Subpopulation structure explained 4% of the phenotypic variance (p-value < 0.0001).



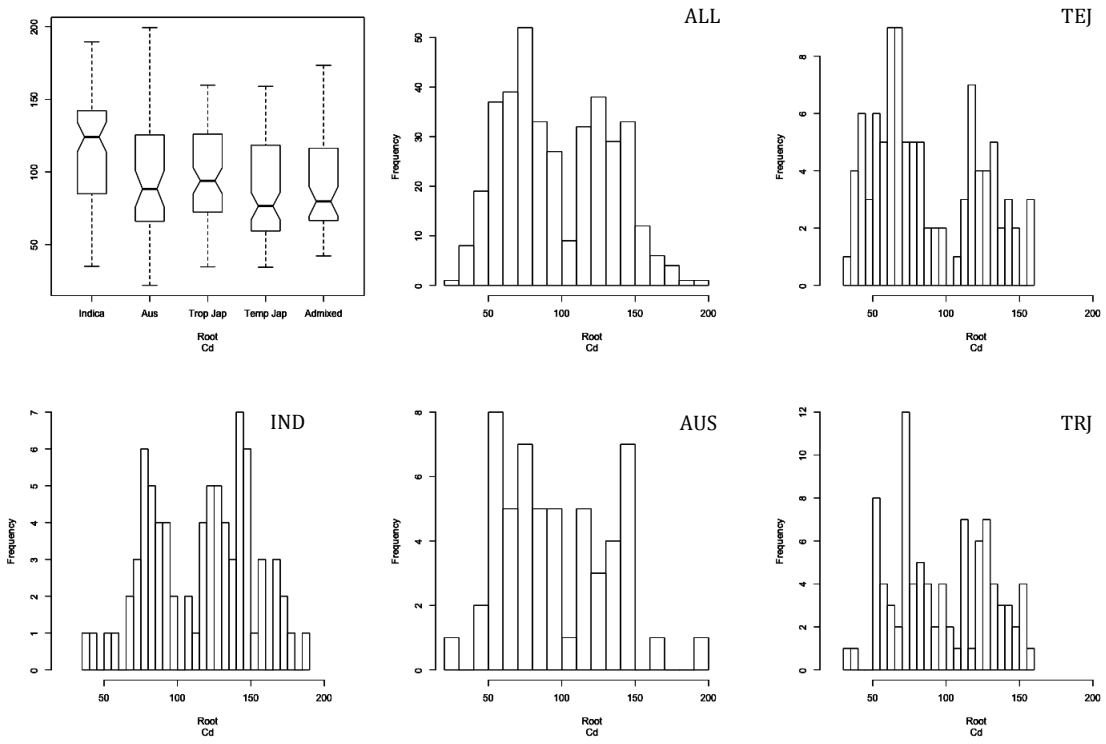
Supplementary Figure 3.56 Boxplot and histogram distributions of boron (B) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for boron shoot content.



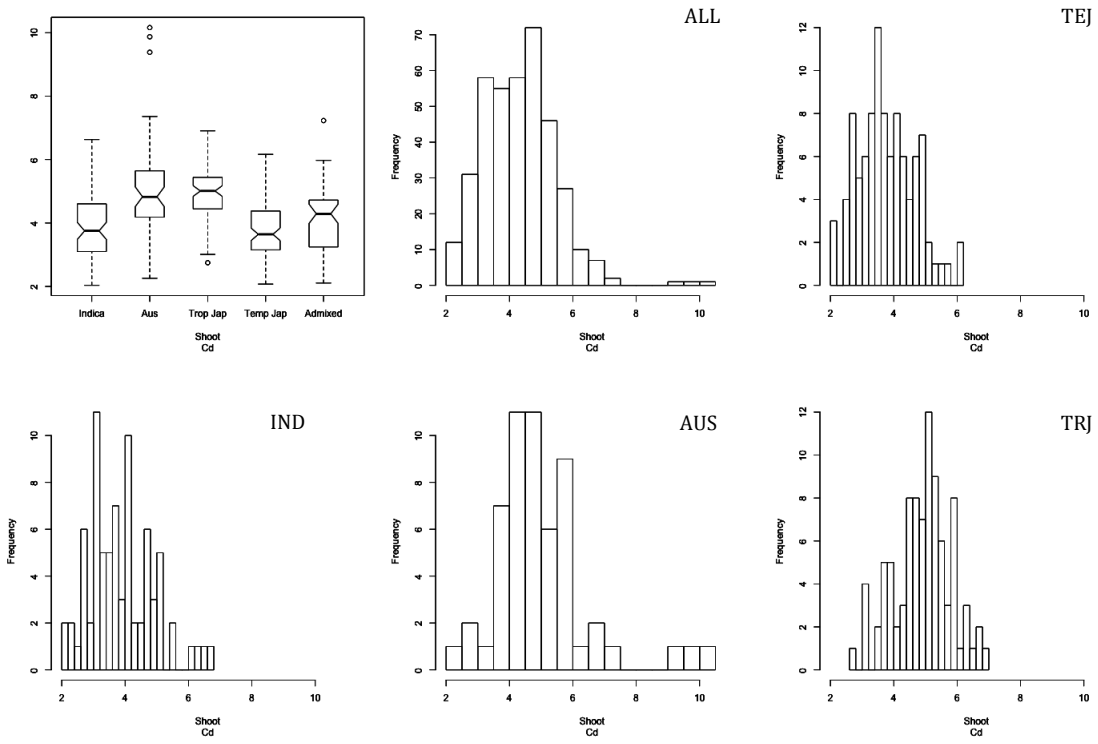
Supplementary Figure 3.57 Boxplot and histogram distributions of calcium (Ca) root phenotype data by subpopulation. Subpopulation structure explained 1% of the phenotypic variance (p-value = .00098).



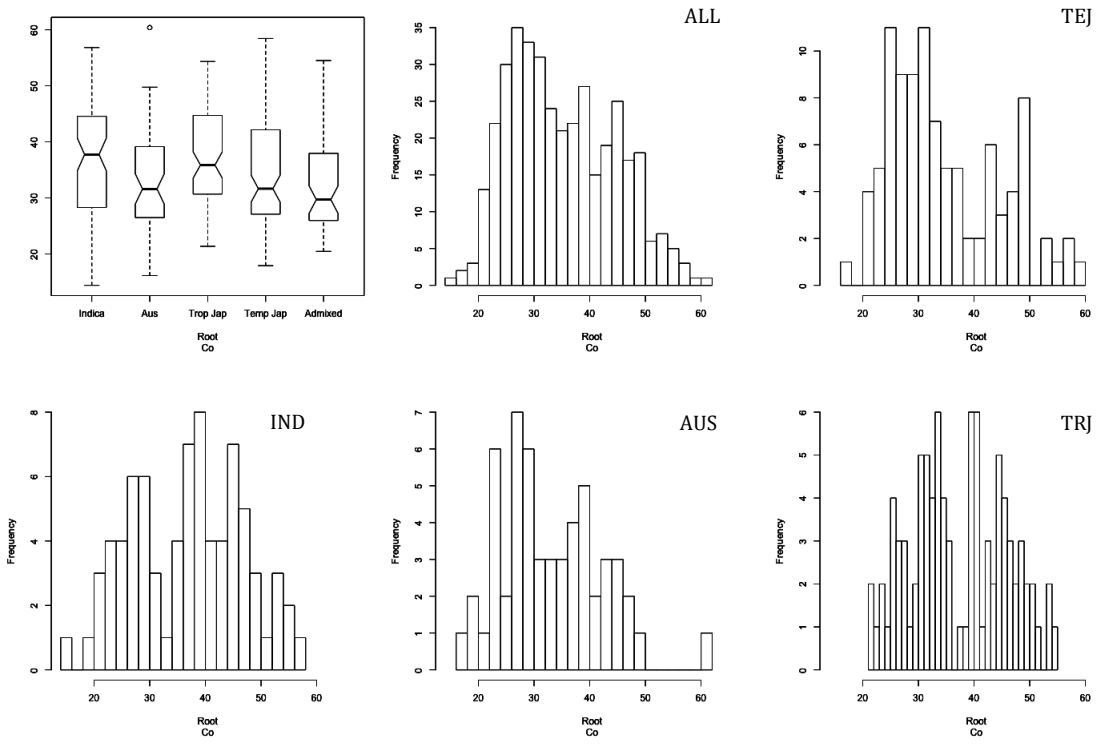
Supplementary Figure 3.58 Boxplot and histogram distributions of calcium (Ca) shoot phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = .00078).



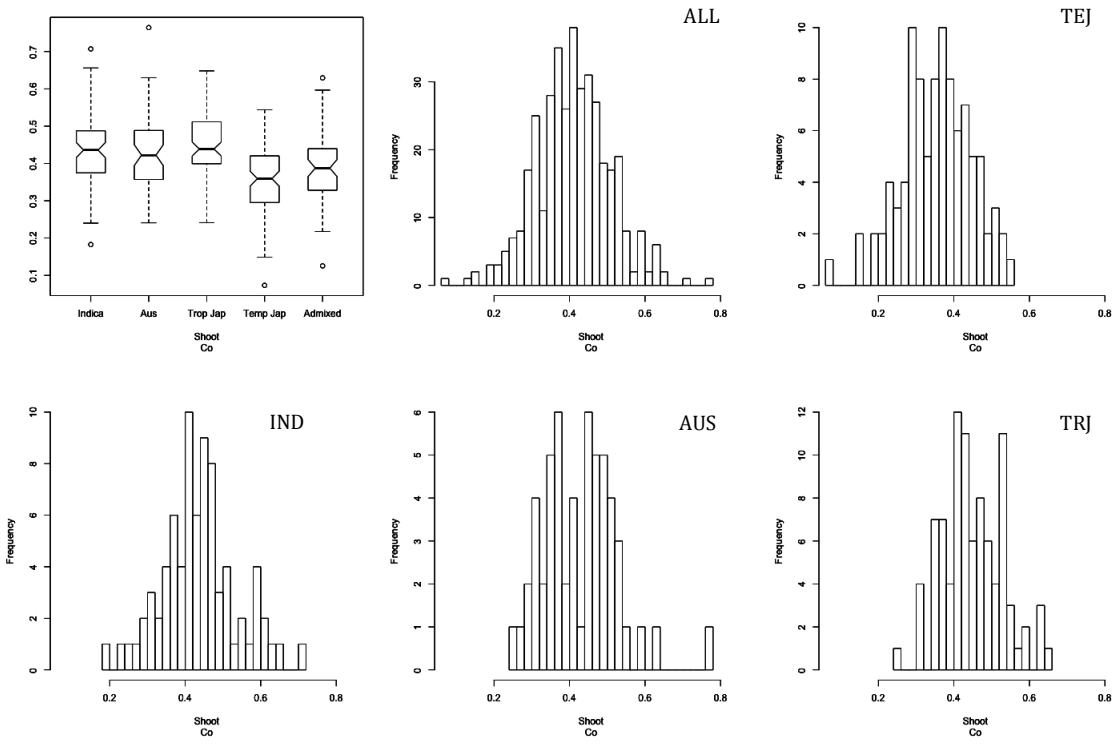
Supplementary Figure 3.59 Boxplot and histogram distributions of cadmium (Cd) root phenotype data by subpopulation. Subpopulation structure explained 6% of the phenotypic variance (p-value < 0.0001).



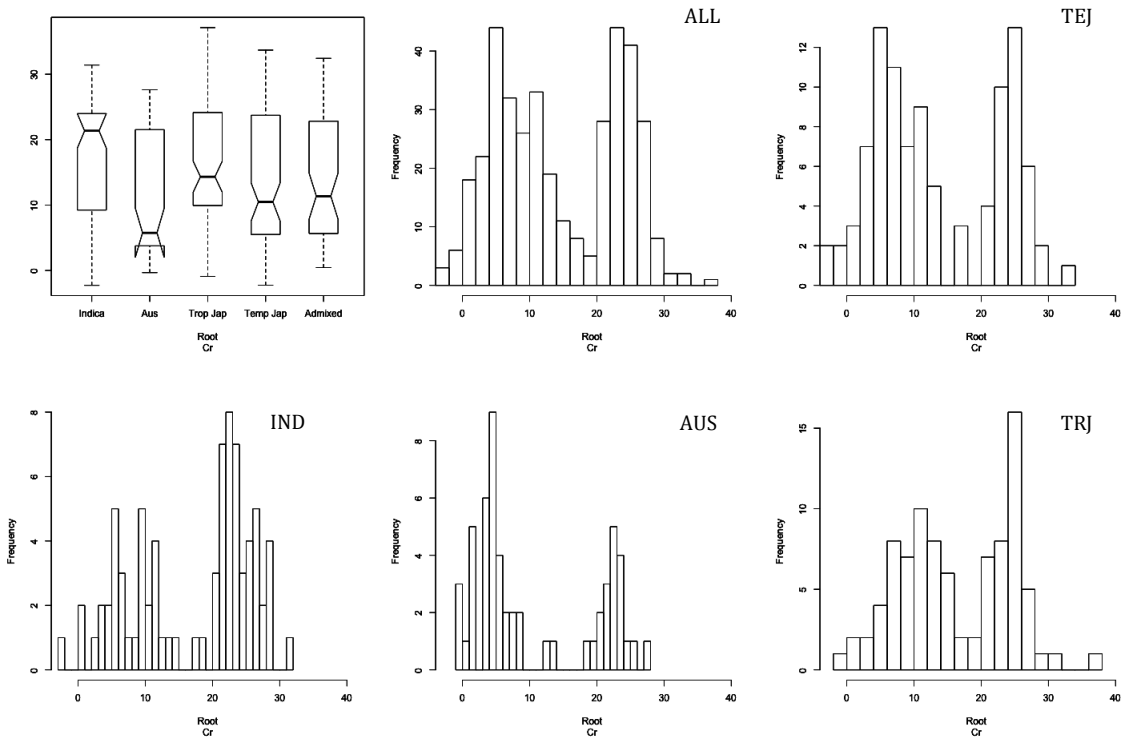
Supplementary Figure 3.60 Boxplot and histogram distributions of cadmium (Cd) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for Cd shoot content.



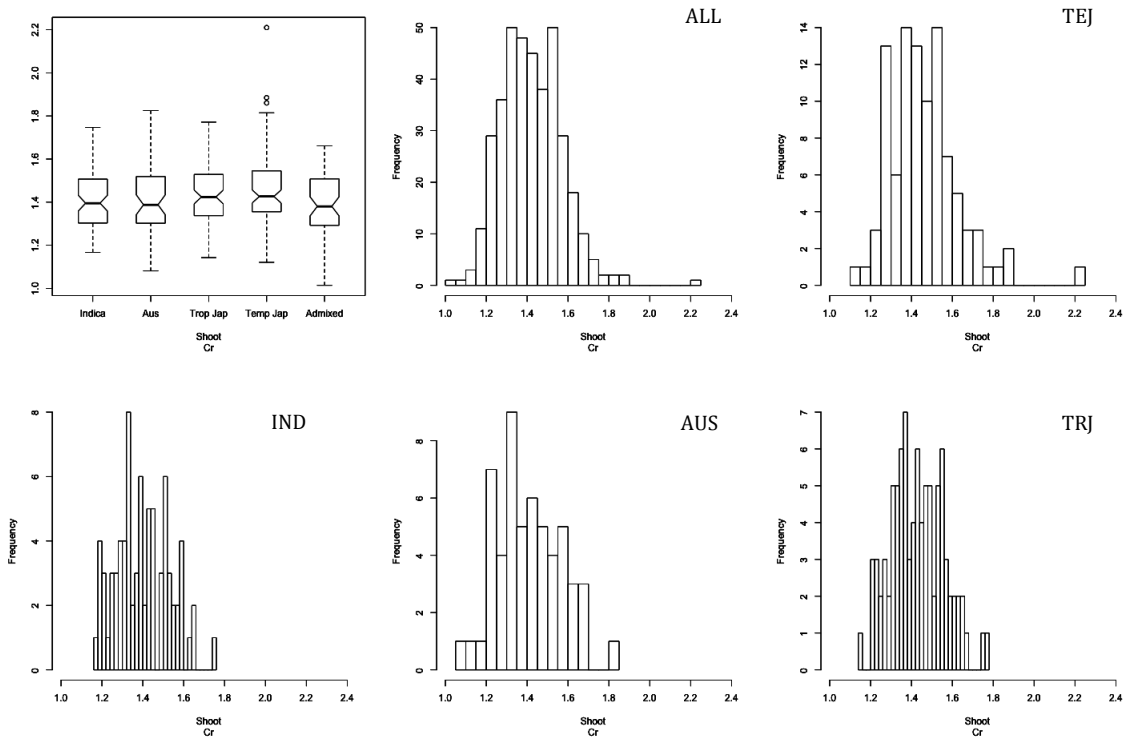
Supplementary Figure 3.61 Boxplot and histogram distributions of cobalt (Co) root phenotype data by subpopulation. Subpopulation structure explained 1% of the phenotypic variance (p-value = 0.0373).



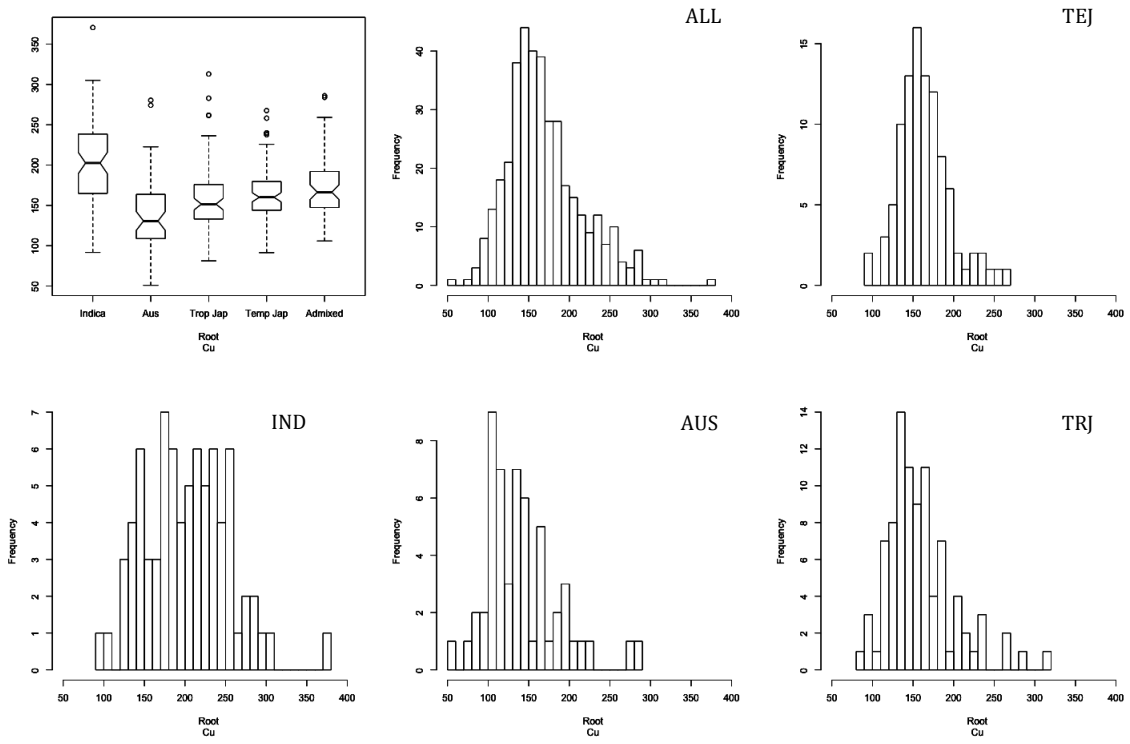
Supplementary Figure 3.62 Boxplot and histogram distributions of cobalt (Co) shoot phenotype data by subpopulation. Subpopulation structure explained 6% of the phenotypic variance (p-value < 0.0001).



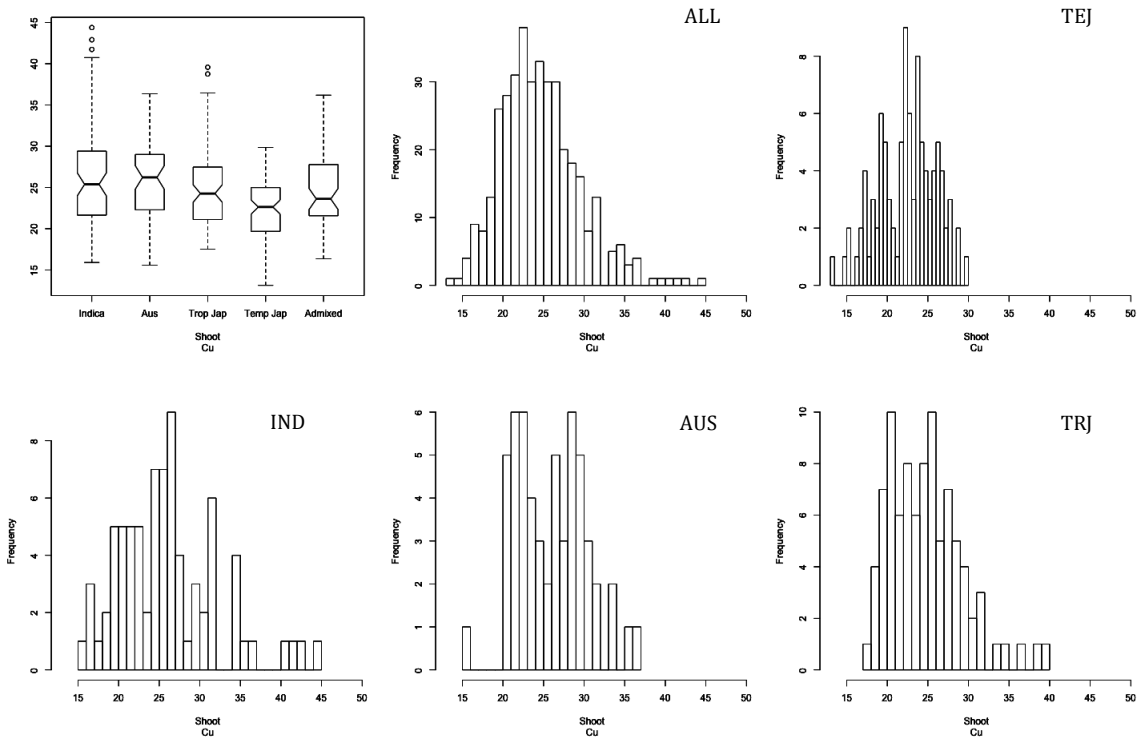
Supplementary Figure 3.63 Boxplot and histogram distributions of chromium (Cr) root phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for Cr root content.



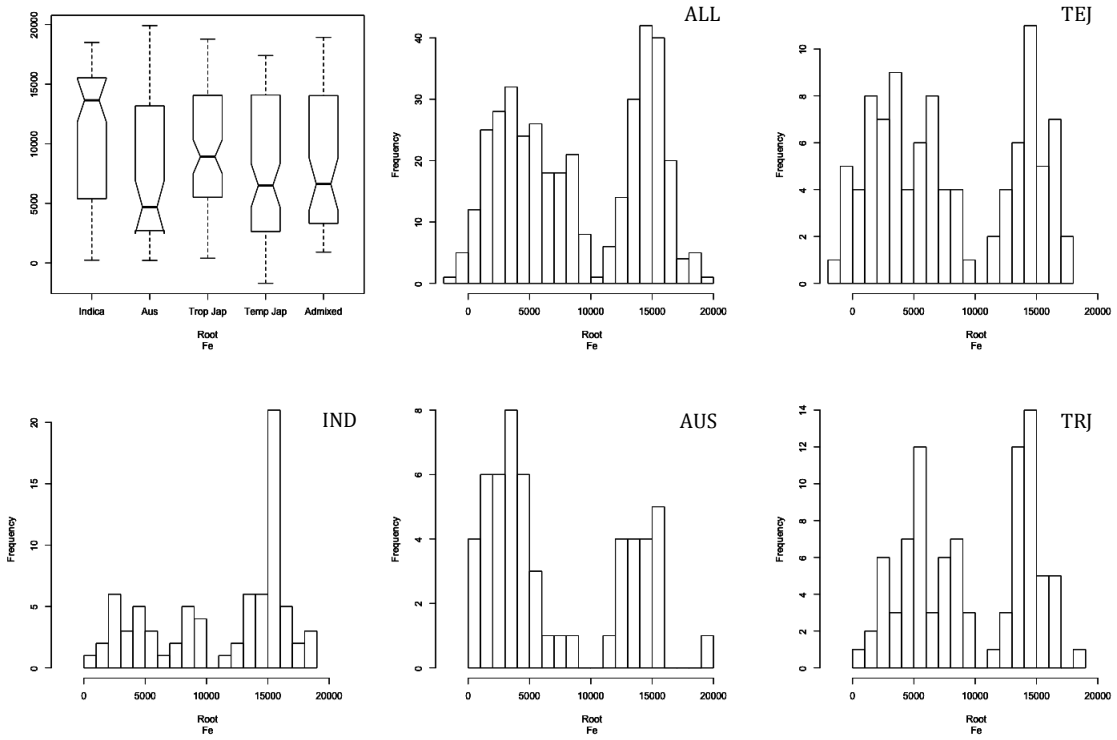
Supplementary Figure 3.64 Boxplot and histogram distributions of chromium (Cr) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for Cr shoot content.



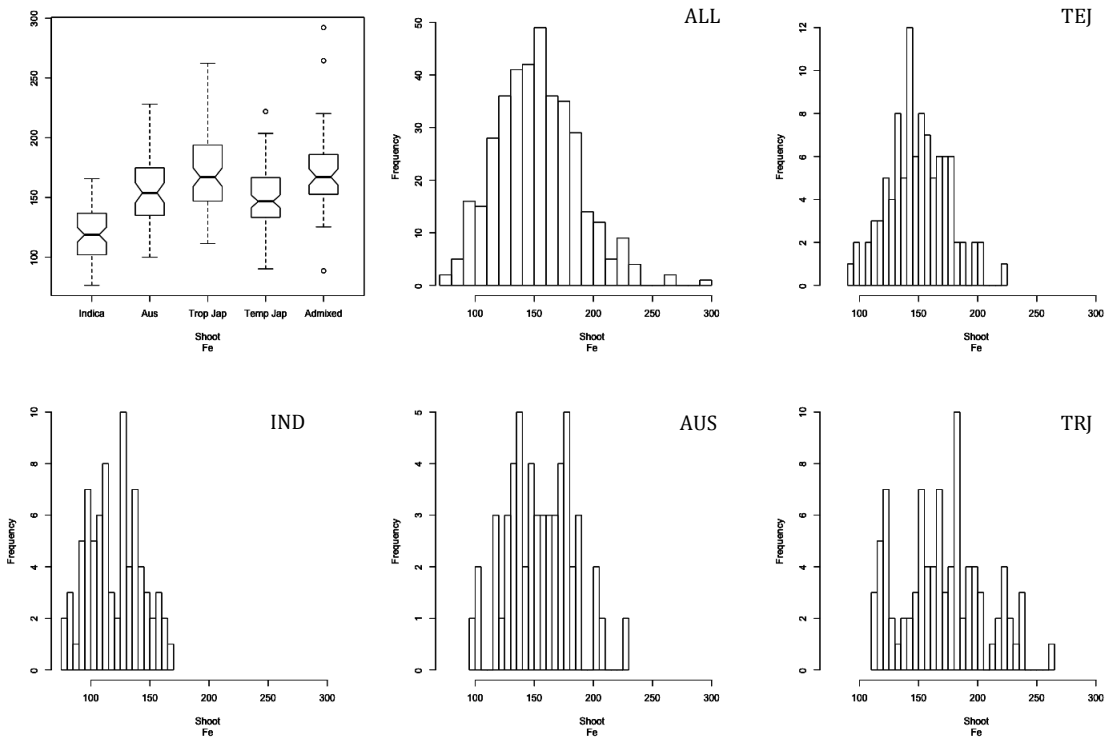
Supplementary Figure 3.65 Boxplot and histogram distributions of copper (Cu) root phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = 0.0052).



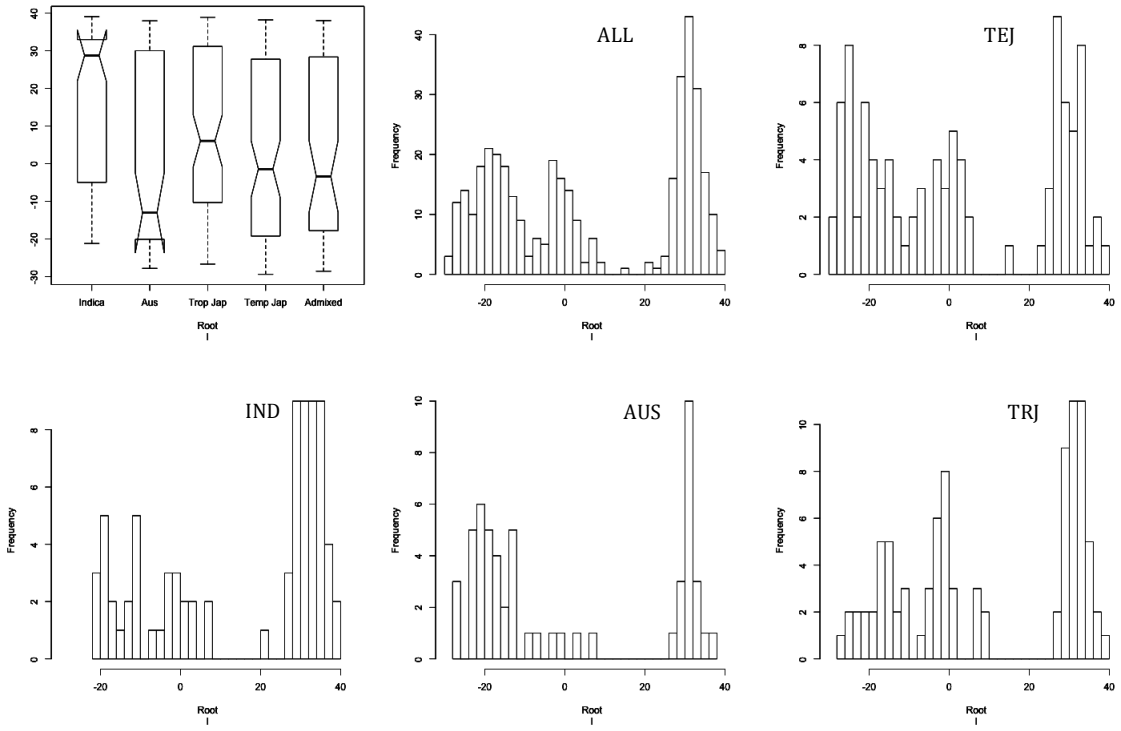
Supplementary Figure 3.66 Boxplot and histogram distributions of copper (Cu) shoot phenotype data by subpopulation. Subpopulation structure explained 4% of the phenotypic variance (p-value < 0.0001).



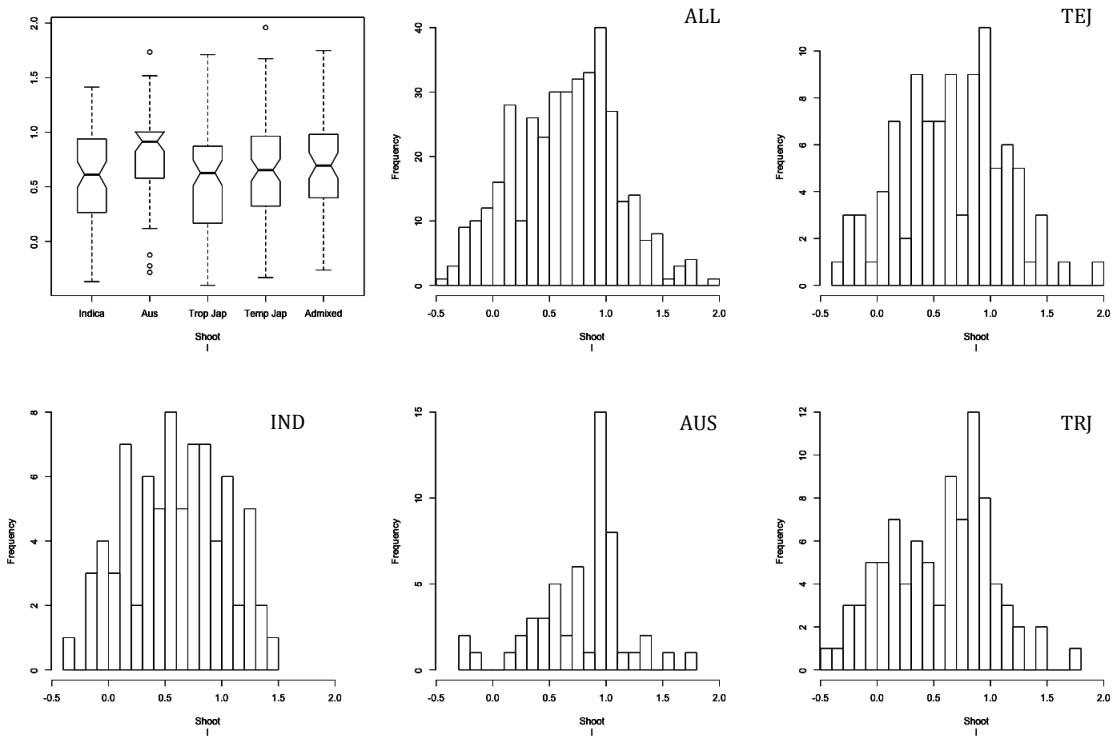
Supplementary Figure 3.67 Boxplot and histogram distributions of iron (Fe) root phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = 0.0013).



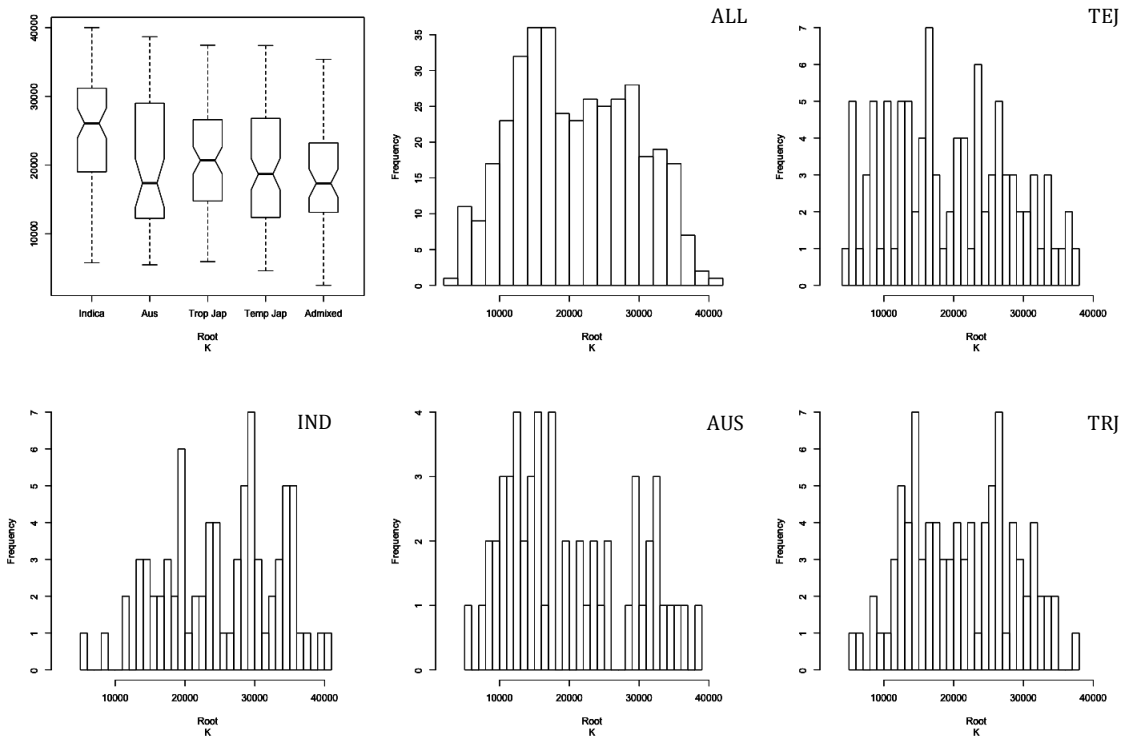
Supplementary Figure 3.68 Boxplot and histogram distributions of iron (Fe) shoot phenotype data by subpopulation. Subpopulation structure explained 1% of the phenotypic variance (p-value < 0.0001).



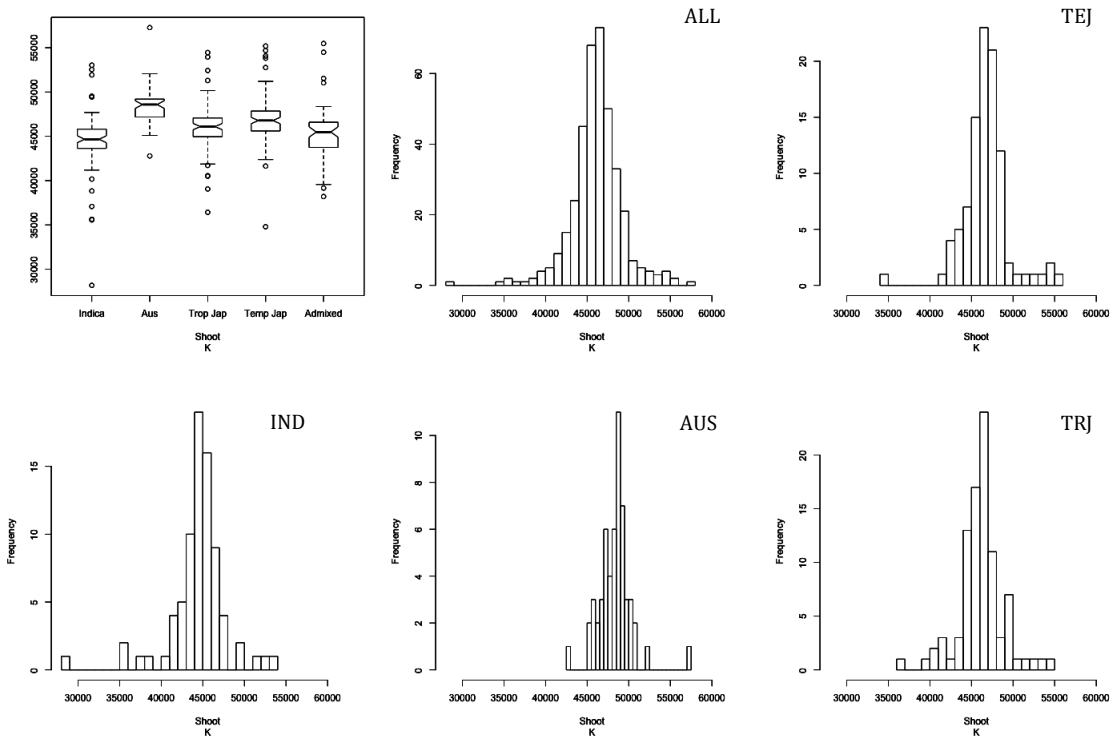
Supplementary Figure 3.69 Boxplot and histogram distributions of iodine (I) root phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0003).



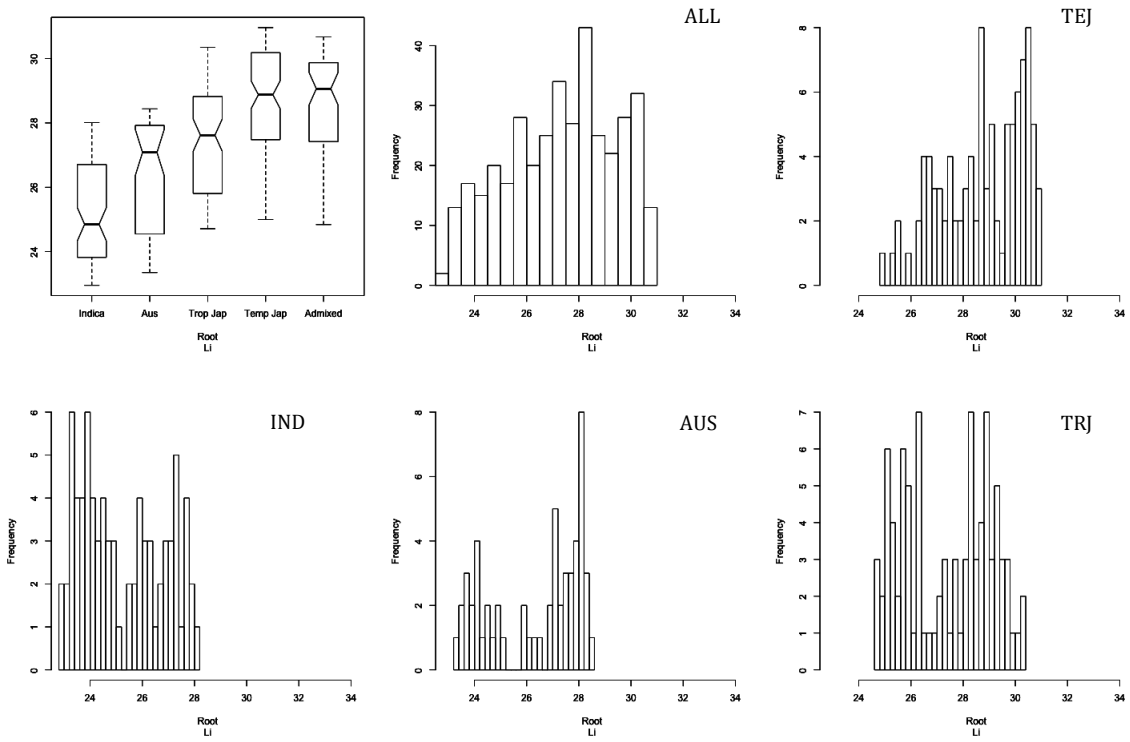
Supplementary Figure 3.70 Boxplot and histogram distributions of iodine (I) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for I shoot content.



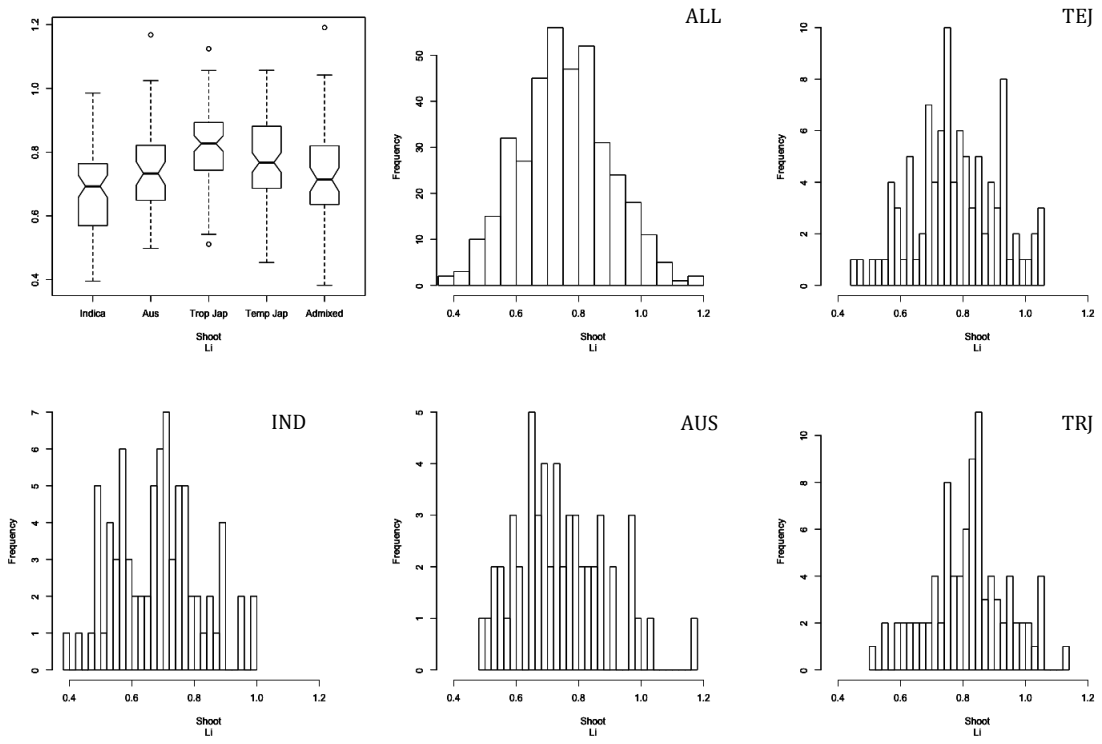
Supplementary Figure 3.71 Boxplot and histogram distributions of potassium (K) root phenotype data by subpopulation. Subpopulation structure explained 6% of the phenotypic variance (p-value < 0.0001).



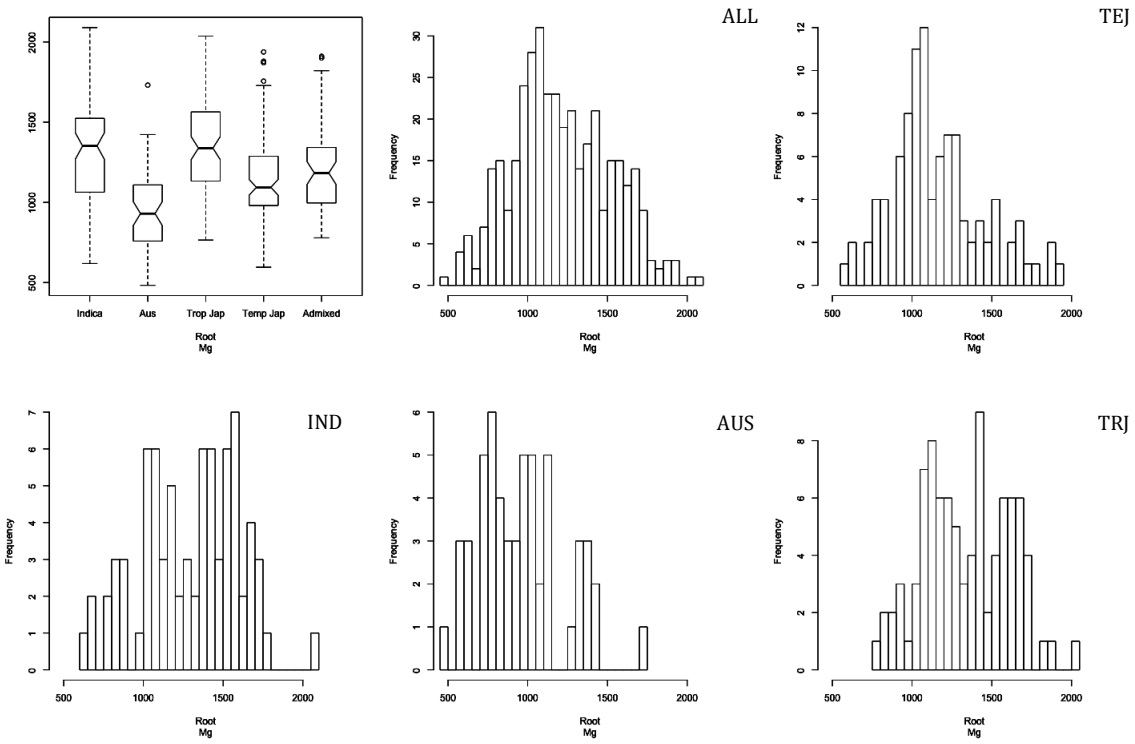
Supplementary Figure 3.72 Boxplot and histogram distributions of potassium (K) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for K shoot content.



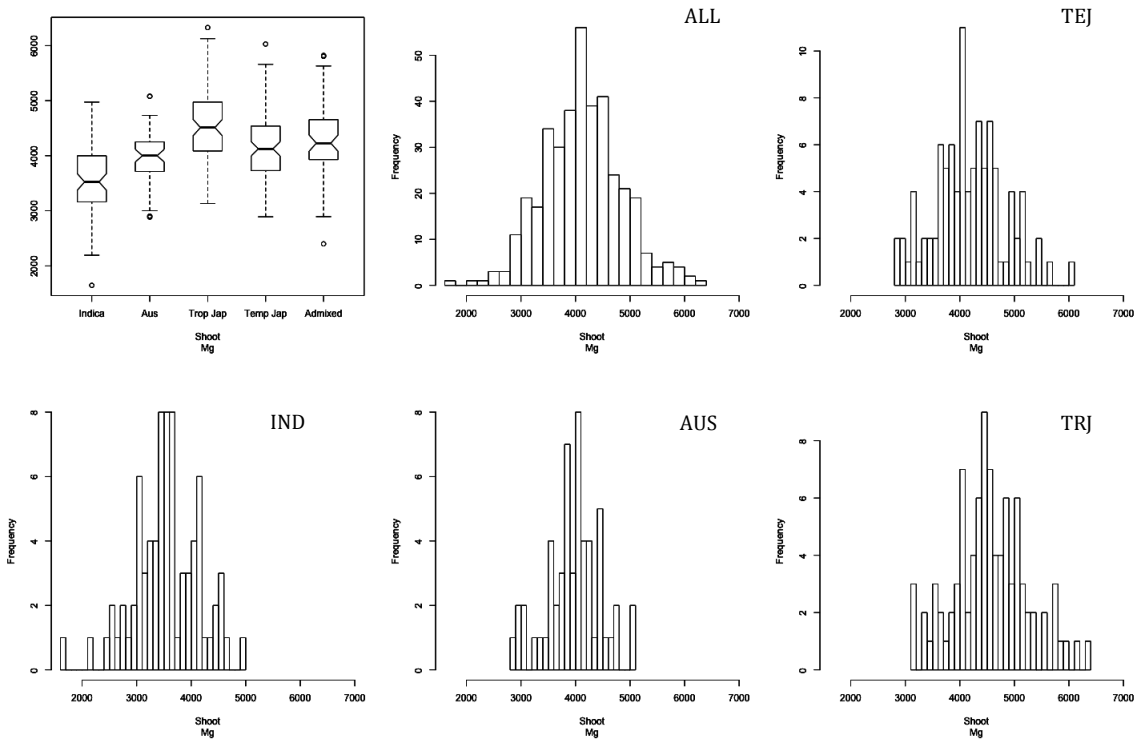
Supplementary Figure 3.73 Boxplot and histogram distributions of lithium (Li) root phenotype data by subpopulation. Subpopulation structure explained 37% of the phenotypic variance (p-value < 0.0001).



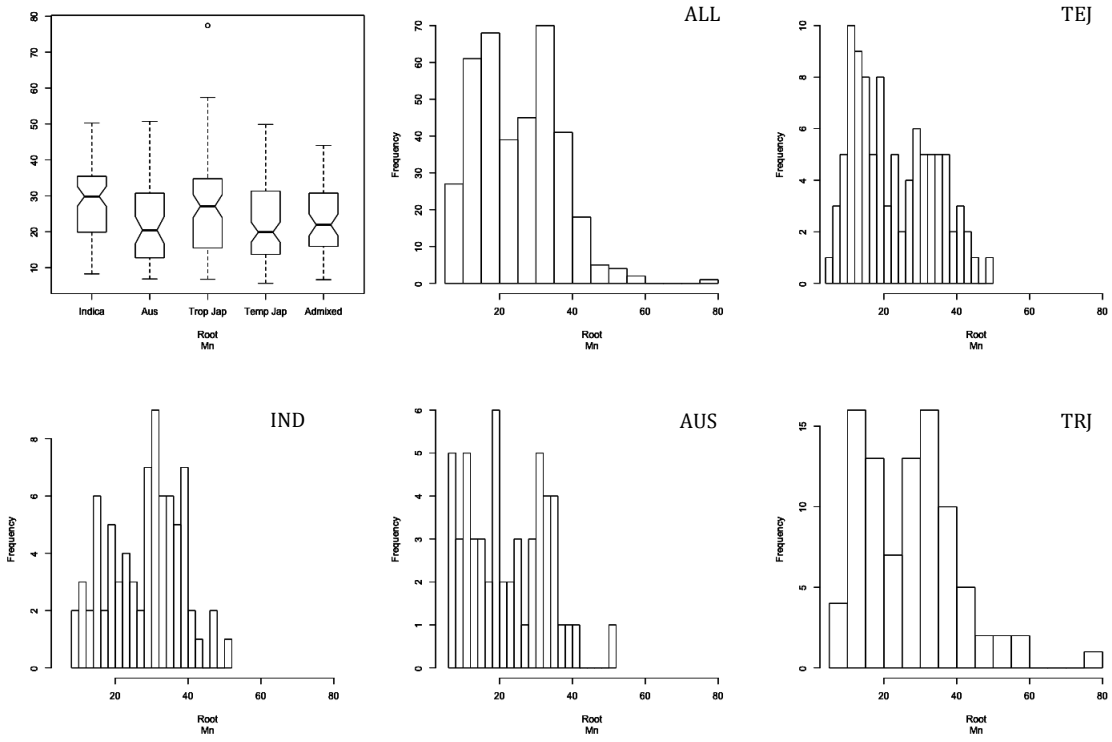
Supplementary Figure 3.74 Boxplot and histogram distributions of lithium (Li) shoot phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = 0.0087).



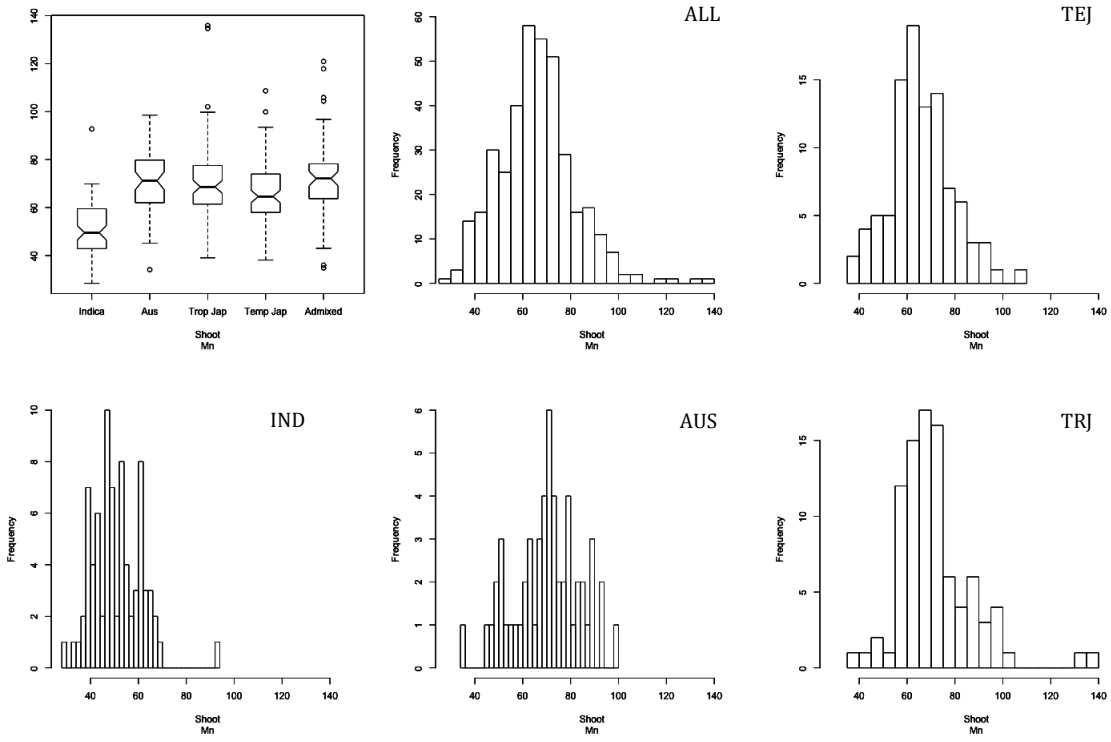
Supplementary Figure 3.75 Boxplot and histogram distributions of magnesium (Mg) root phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for Mg root content.



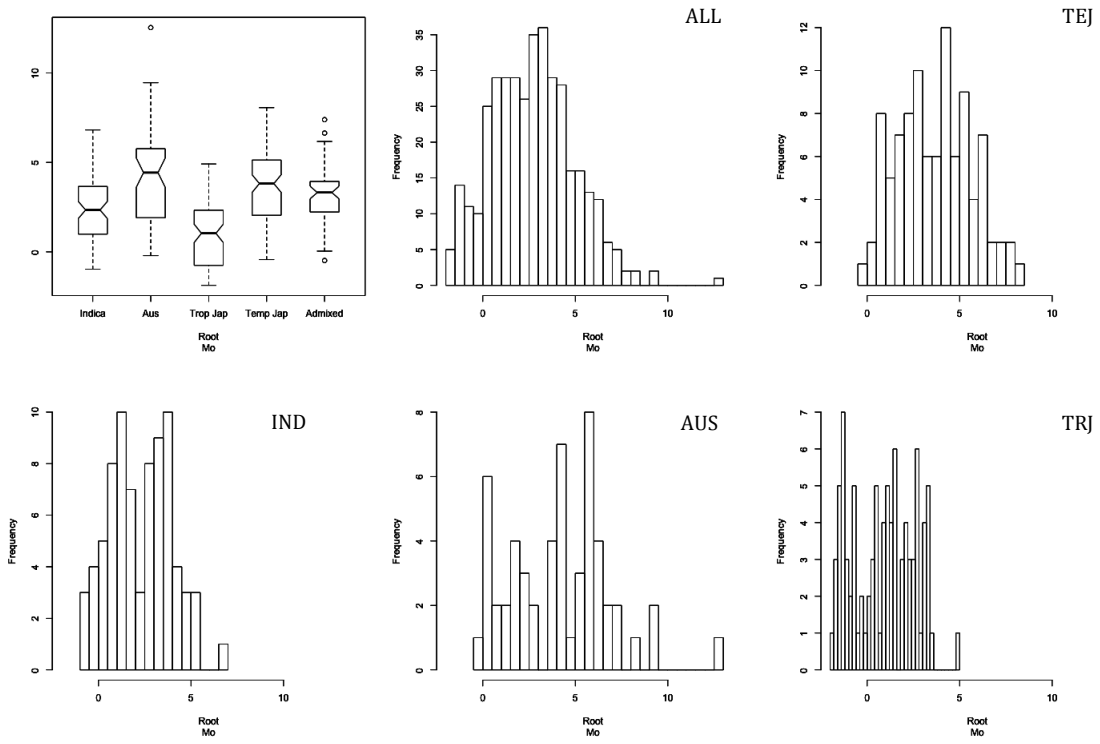
Supplementary Figure 3.76 Boxplot and histogram distributions of magnesium (Mg) shoot phenotype data by subpopulation. Subpopulation structure explained 10% of the phenotypic variance (p-value < 0.0001).



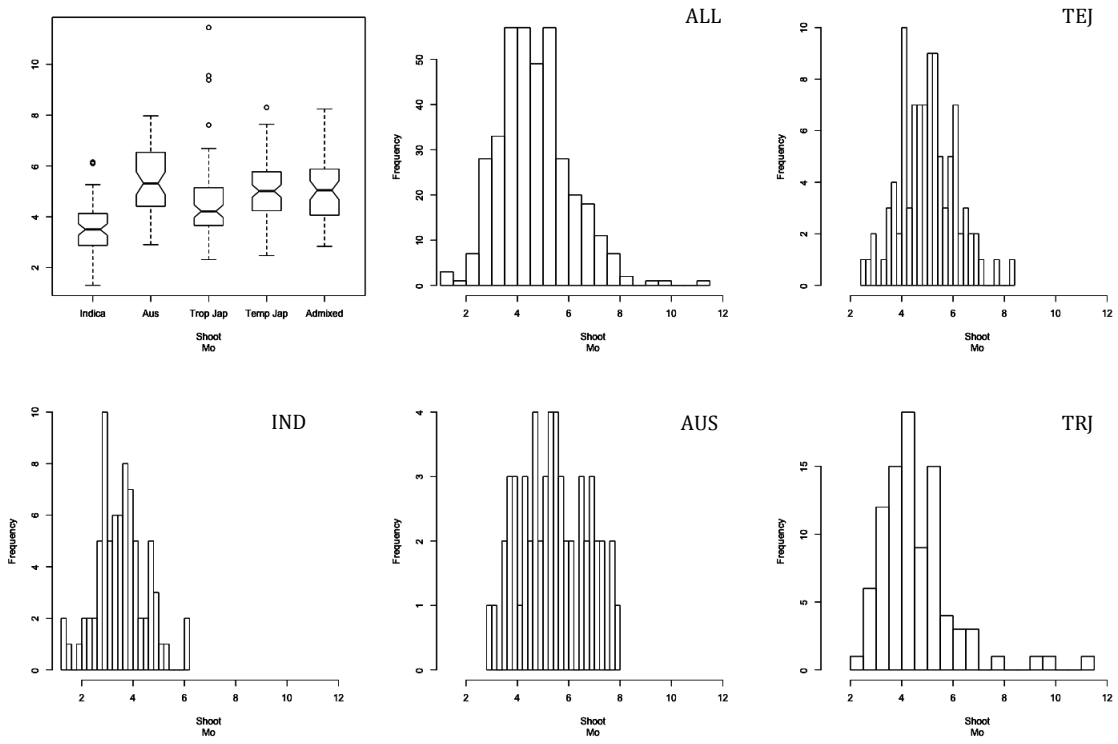
Supplementary Figure 3.77 Boxplot and histogram distributions of manganese (Mn) root phenotype data by subpopulation. Subpopulation structure explained 1% of the phenotypic variance (p-value = 0.0087).



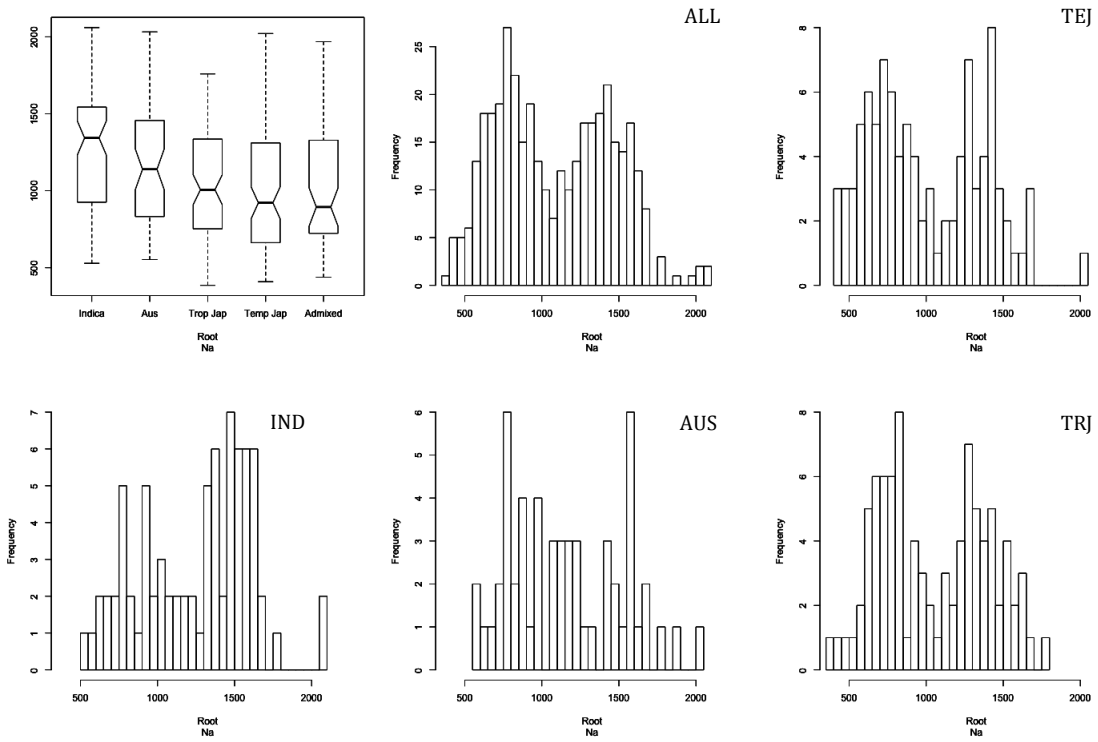
Supplementary Figure 3.78 Boxplot and histogram distributions of manganese (Mn) shoot phenotype data by subpopulation. Subpopulation structure explained 11% of the phenotypic variance (p-value < 0.0001).



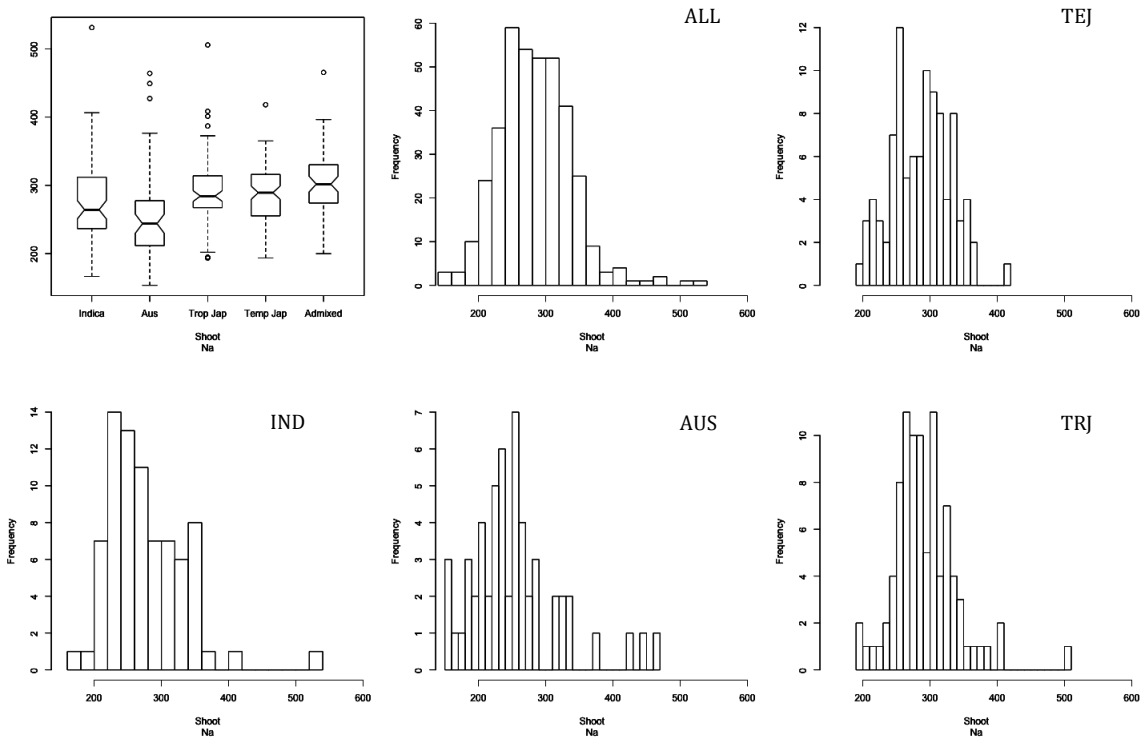
Supplementary Figure 3.79 Boxplot and histogram distributions of molybdenum (Mo) root phenotype data by subpopulation. Subpopulation structure explained 1% of the phenotypic variance (p-value < 0.0362).



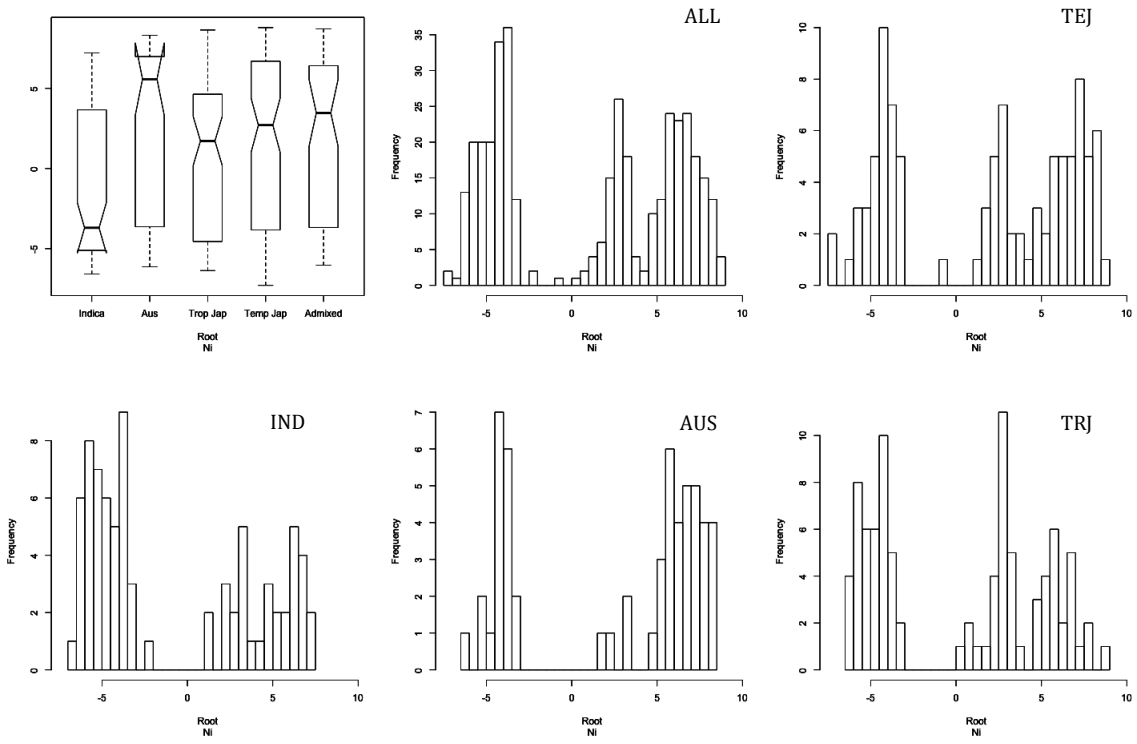
Supplementary Figure 3.80 Boxplot and histogram distributions of molybdenum (Mo) shoot phenotype data by subpopulation. Subpopulation structure explained 10% of the phenotypic variance (p-value < 0.0001).



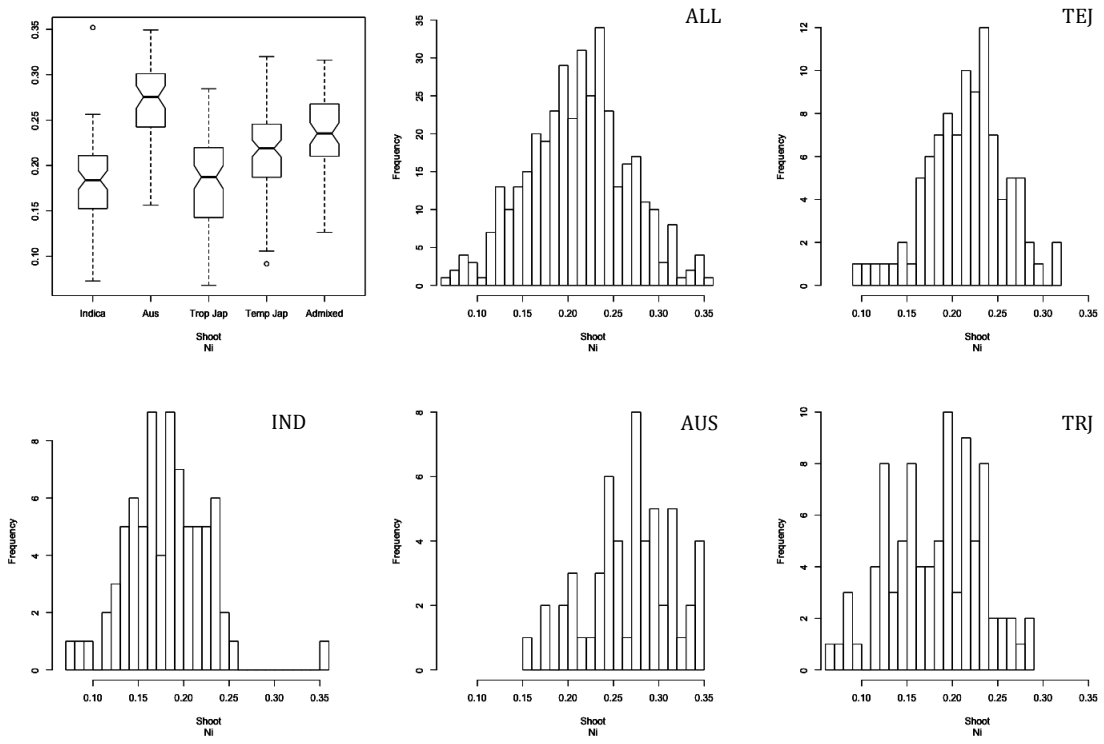
Supplementary Figure 3.81 Boxplot and histogram distributions of sodium (Na) root phenotype data by subpopulation. Subpopulation structure explained 6% of the phenotypic variance (p-value < 0.0001).



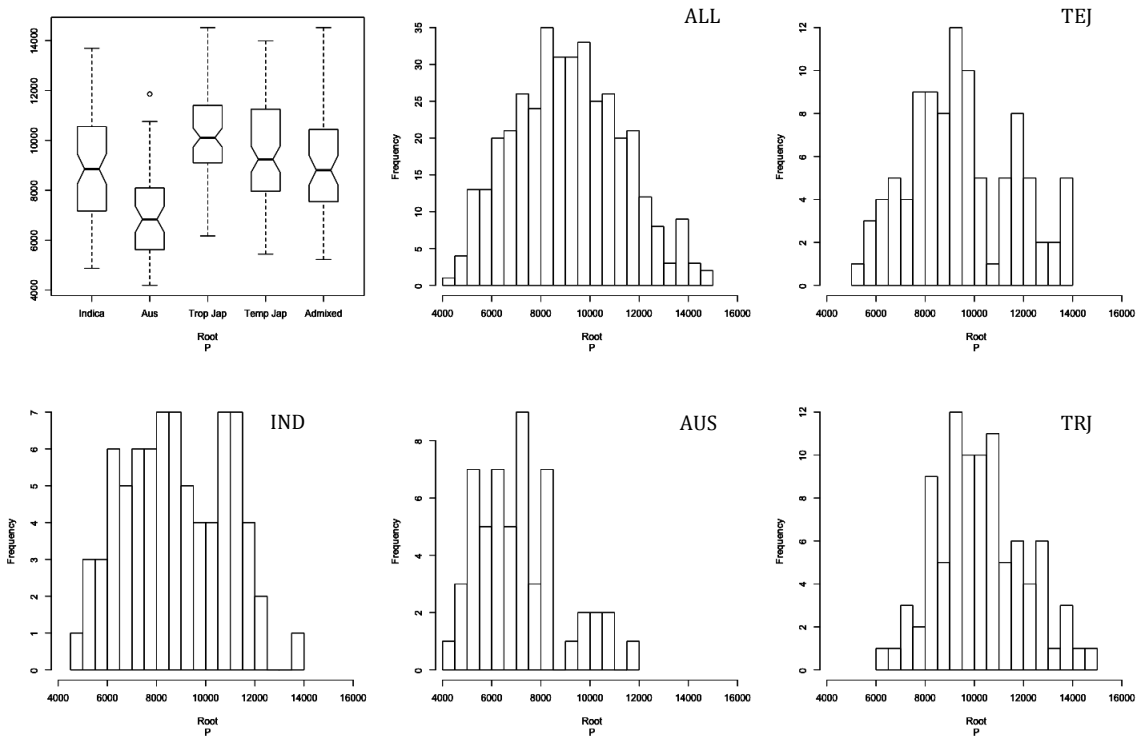
Supplementary Figure 3.82 Boxplot and histogram distributions of sodium (Na) shoot phenotype data by subpopulation. Subpopulation structure explained 4% of the phenotypic variance (p-value < 0.0001).



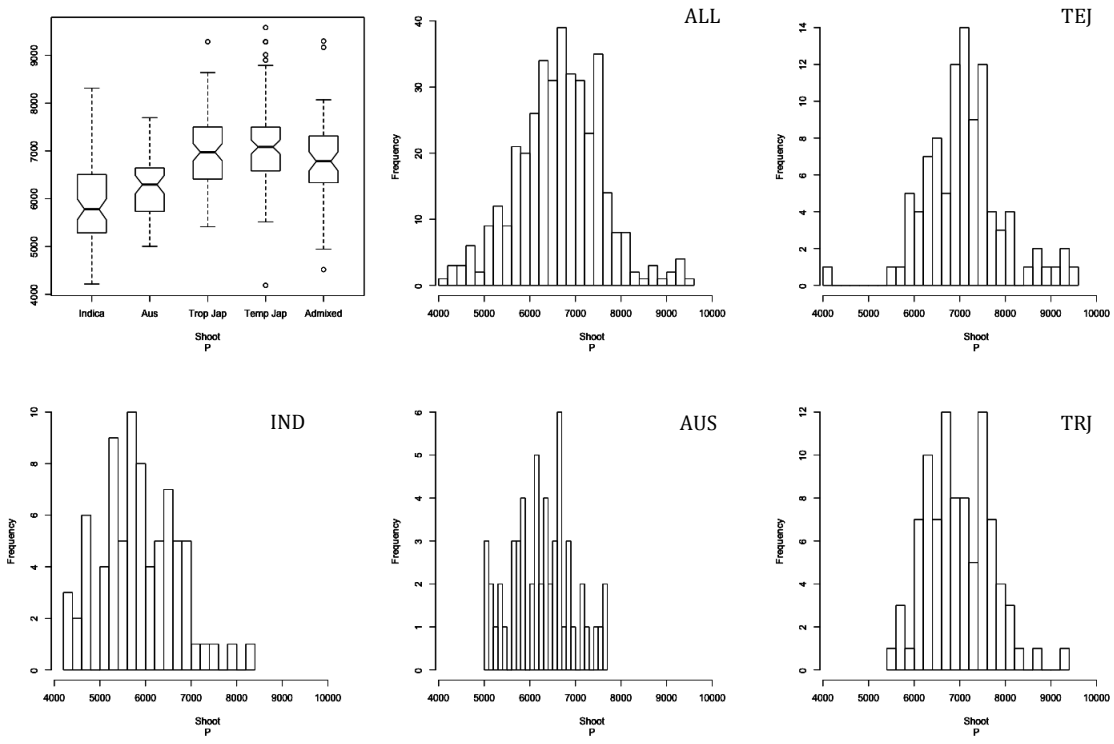
Supplementary Figure 3.83 Boxplot and histogram distributions of nickel (Ni) root phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0006).



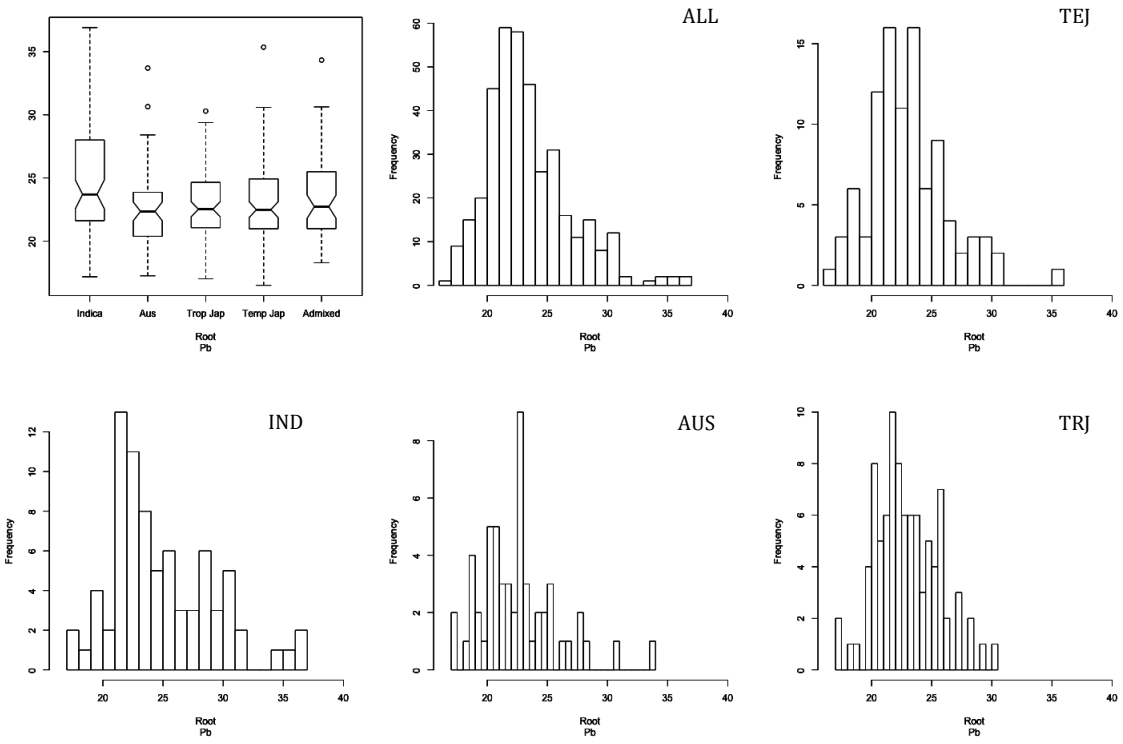
Supplementary Figure 3.84 Boxplot and histogram distributions of nickel (Ni) shoot phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.001).



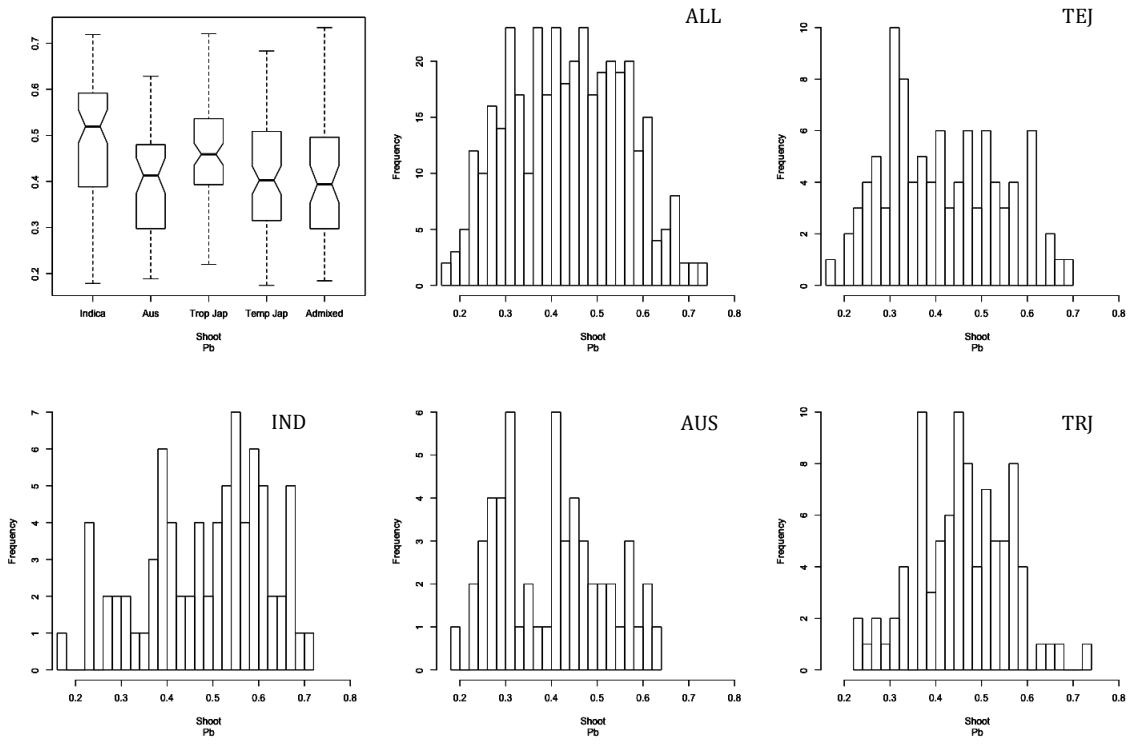
Supplementary Figure 3.85 Boxplot and histogram distributions of phosphorus (P) root phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0011).



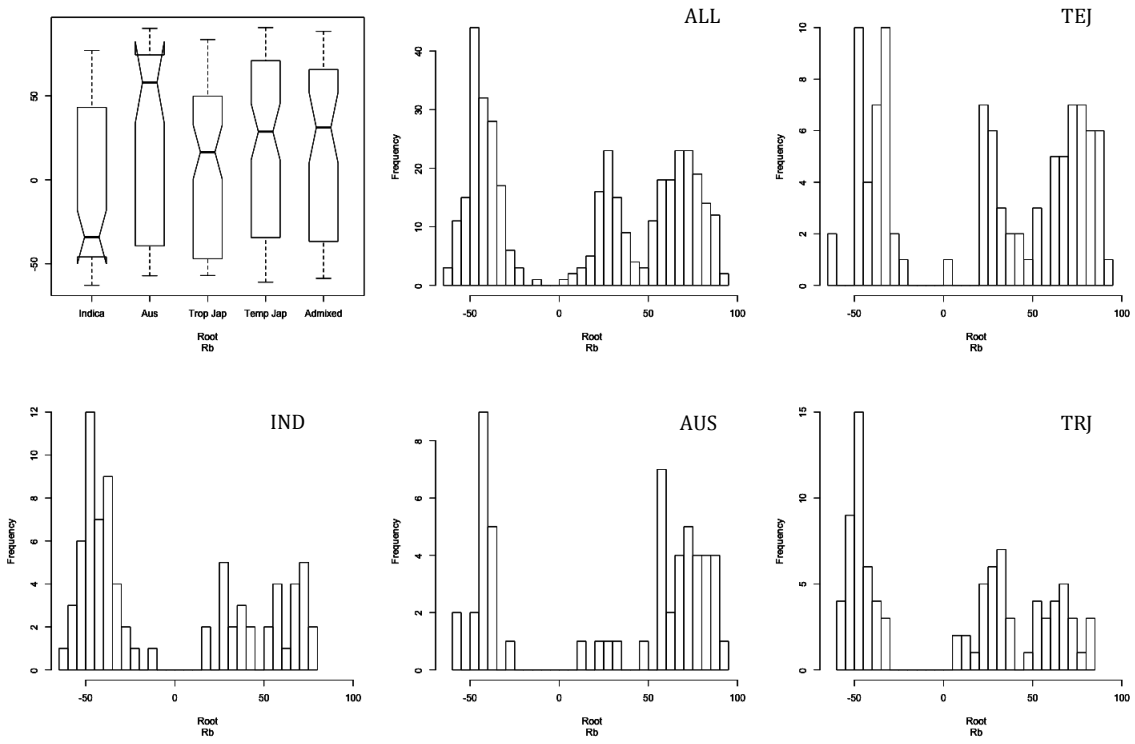
Supplementary Figure 3.86 Boxplot and histogram distributions of phosphorus (P) shoot phenotype data by subpopulation. Subpopulation structure explained 18% of the phenotypic variance (p-value < 0.0001).



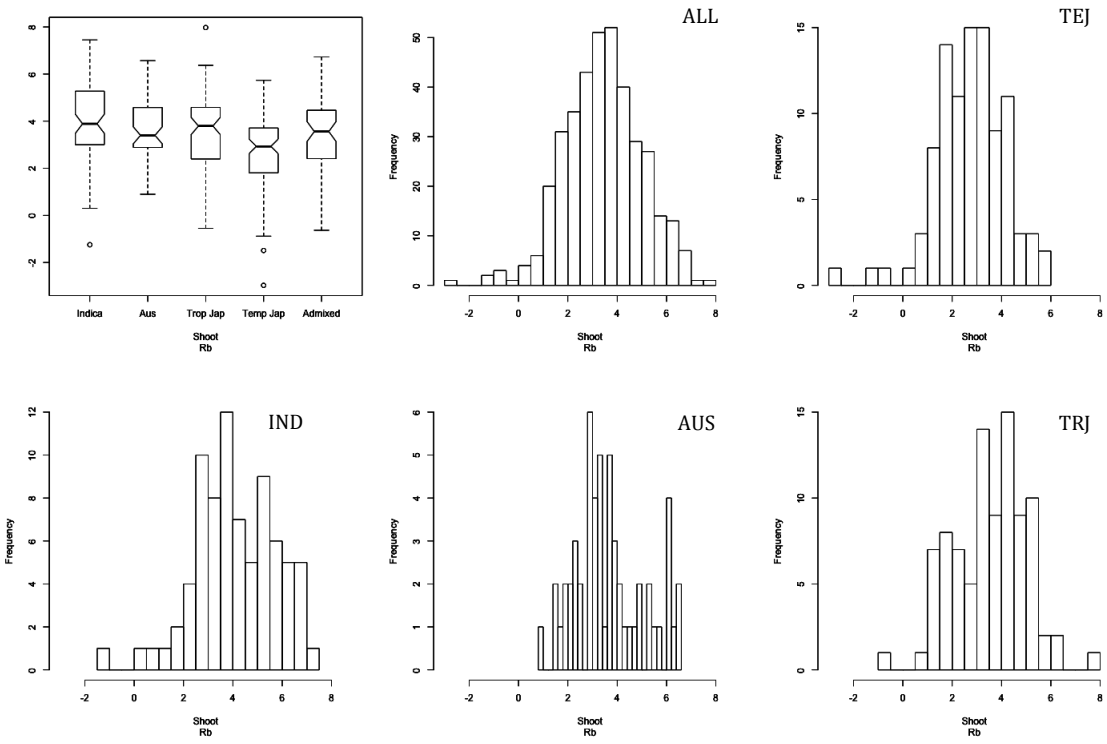
Supplementary Figure 3.87 Boxplot and histogram distributions of lead (Pb) root phenotype data by subpopulation. Subpopulation structure explained 1% of the phenotypic variance (p-value = 0.0198).



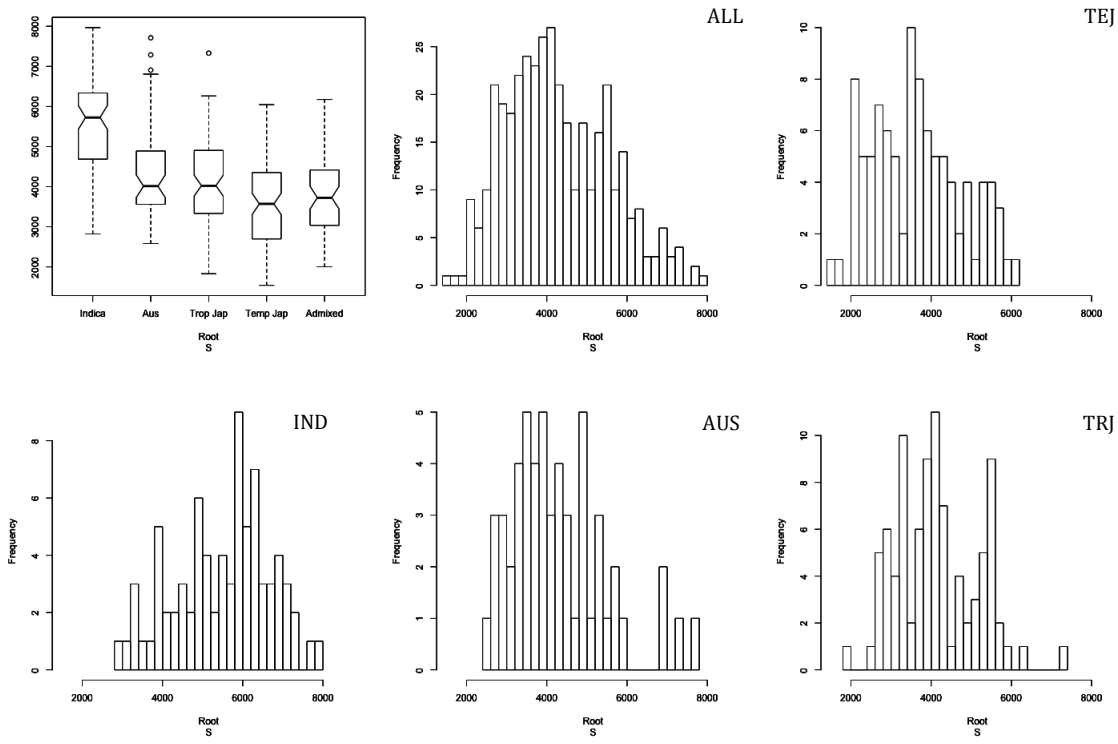
Supplementary Figure 3.88 Boxplot and histogram distributions of lead (Pb) shoot phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0004).



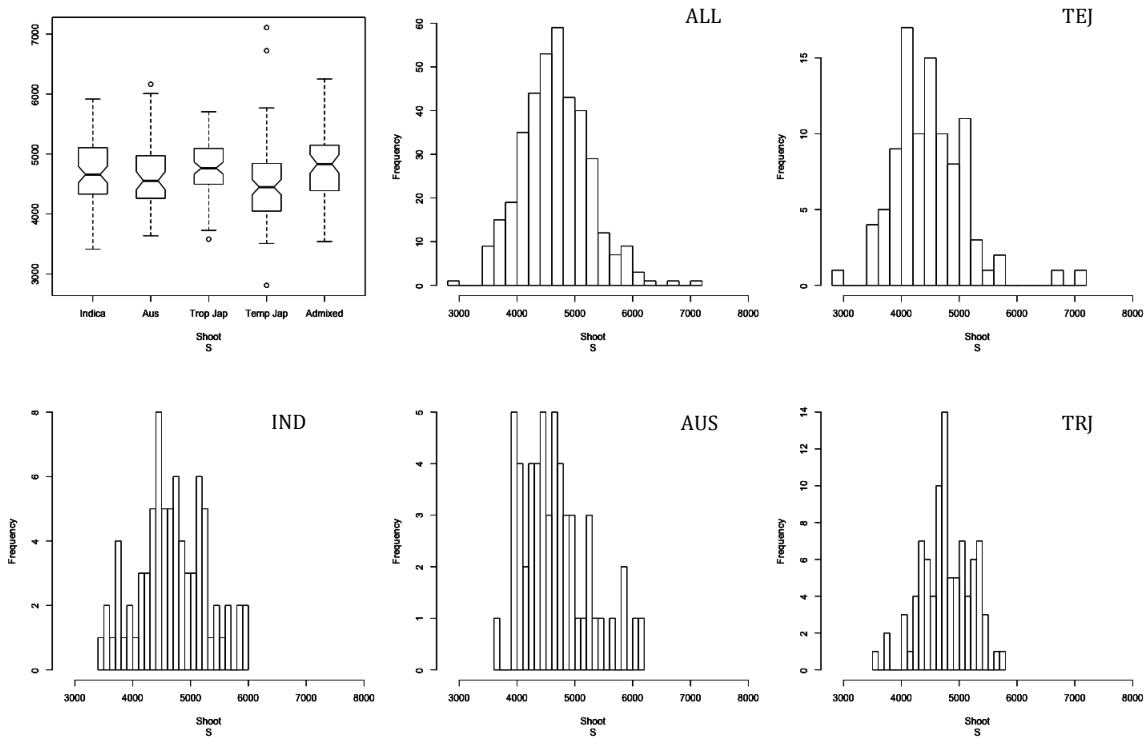
Supplementary Figure 3.89 Boxplot and histogram distributions of rubidium (Rb) root phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = 0.0021).



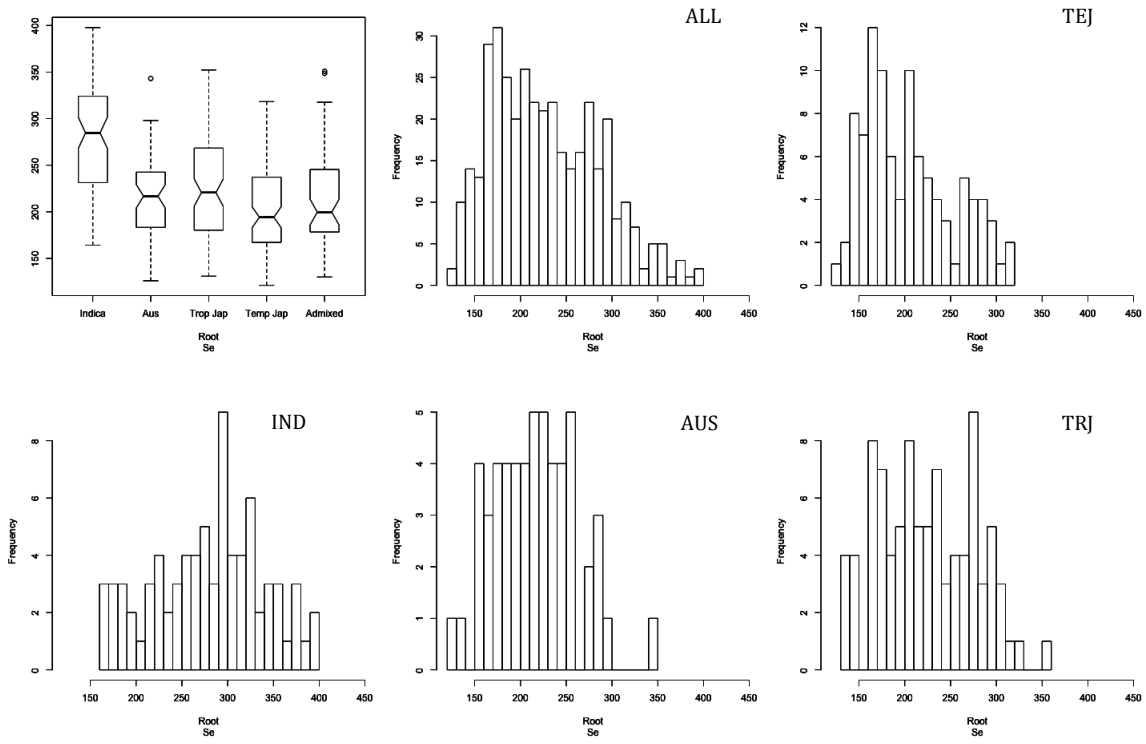
Supplementary Figure 3.90 Boxplot and histogram distributions of rubidium (Rb) shoot phenotype data by subpopulation. Subpopulation structure explained 5% of the phenotypic variance (p-value < 0.0001).



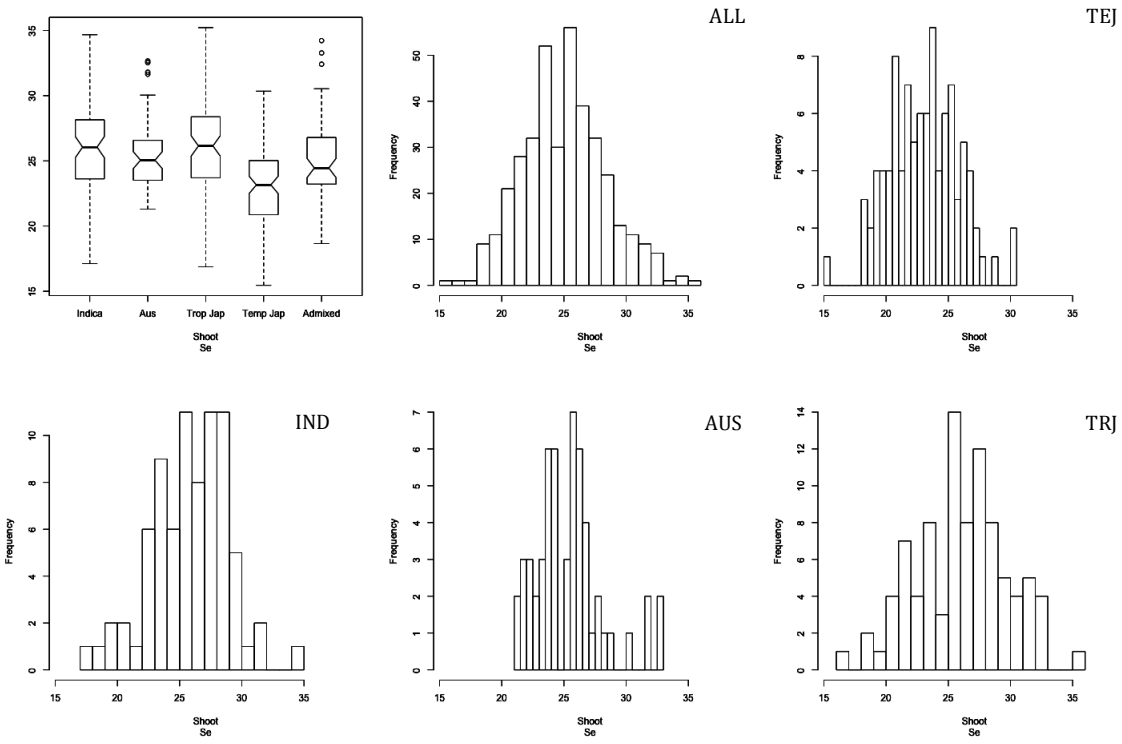
Supplementary Figure 3.91 Boxplot and histogram distributions of sulfur (S) root phenotype data by subpopulation. Subpopulation structure explained 20% of the phenotypic variance (p-value < 0.0001).



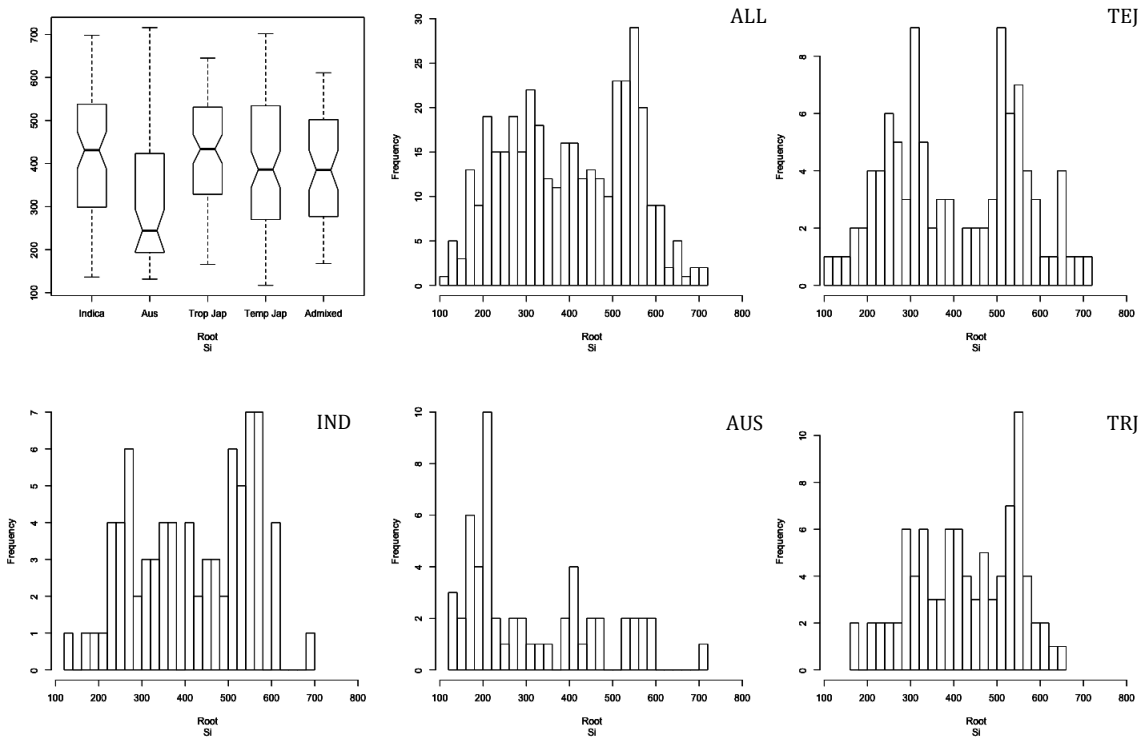
Supplementary Figure 3.92 Boxplot and histogram distributions of sulfur (S) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for S shoot content.



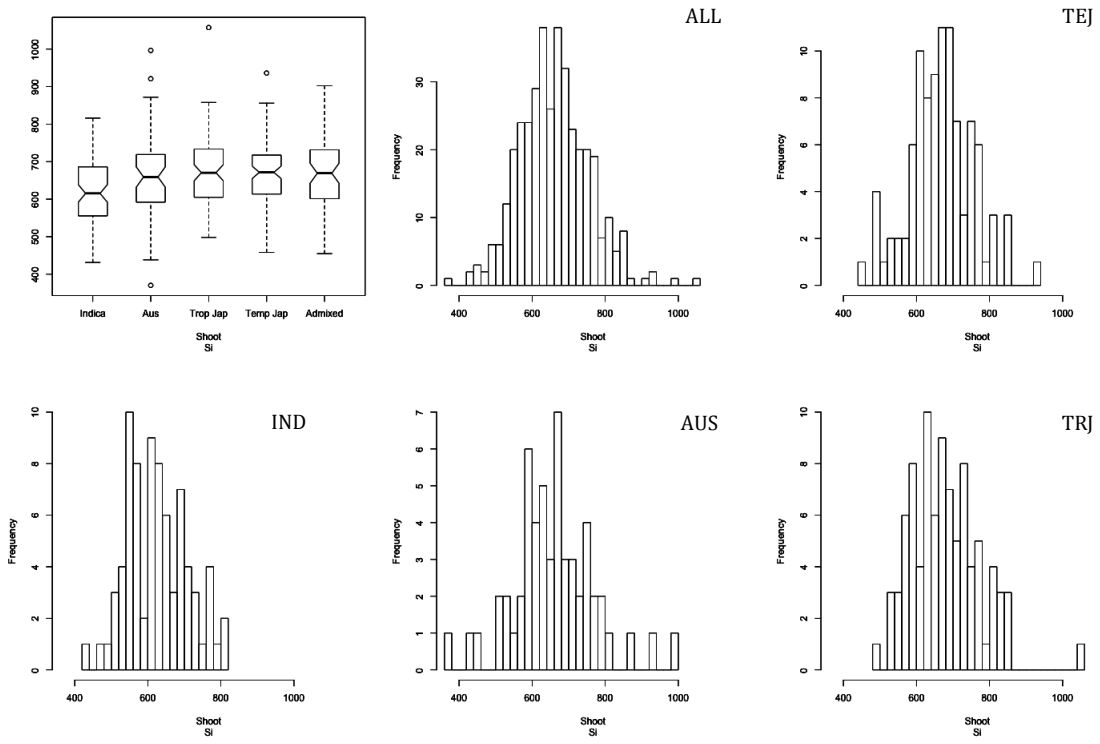
Supplementary Figure 3.93 Boxplot and histogram distributions of selenium (Se) root phenotype data by subpopulation. Subpopulation structure explained 13% of the phenotypic variance (p-value < 0.0001).



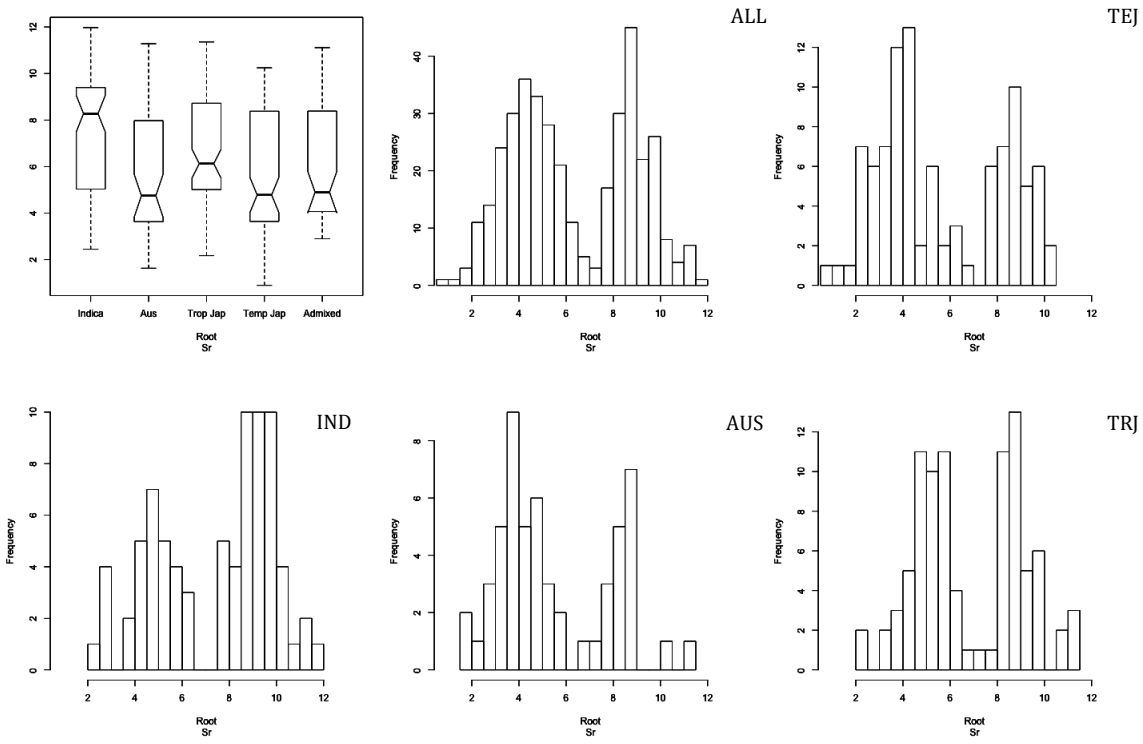
Supplementary Figure 3.94 Boxplot and histogram distributions of selenium (Se) shoot phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0004).



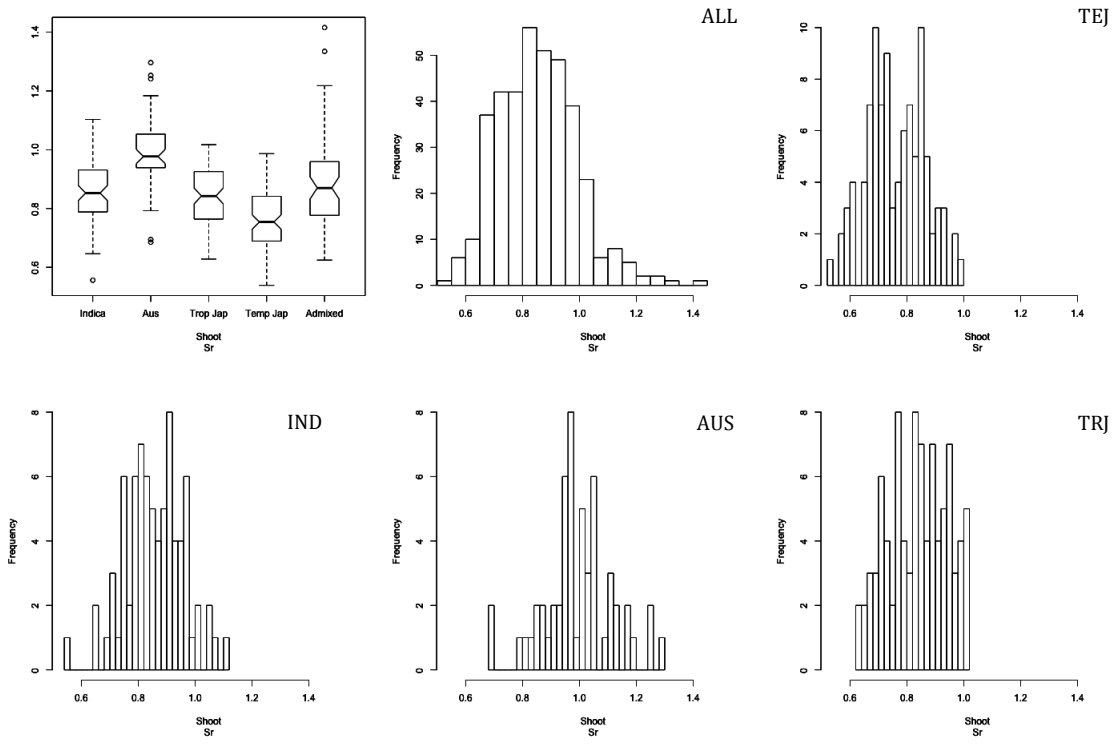
Supplementary Figure 3.95 Boxplot and histogram distributions of silicon (Si) root phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for Si root content.



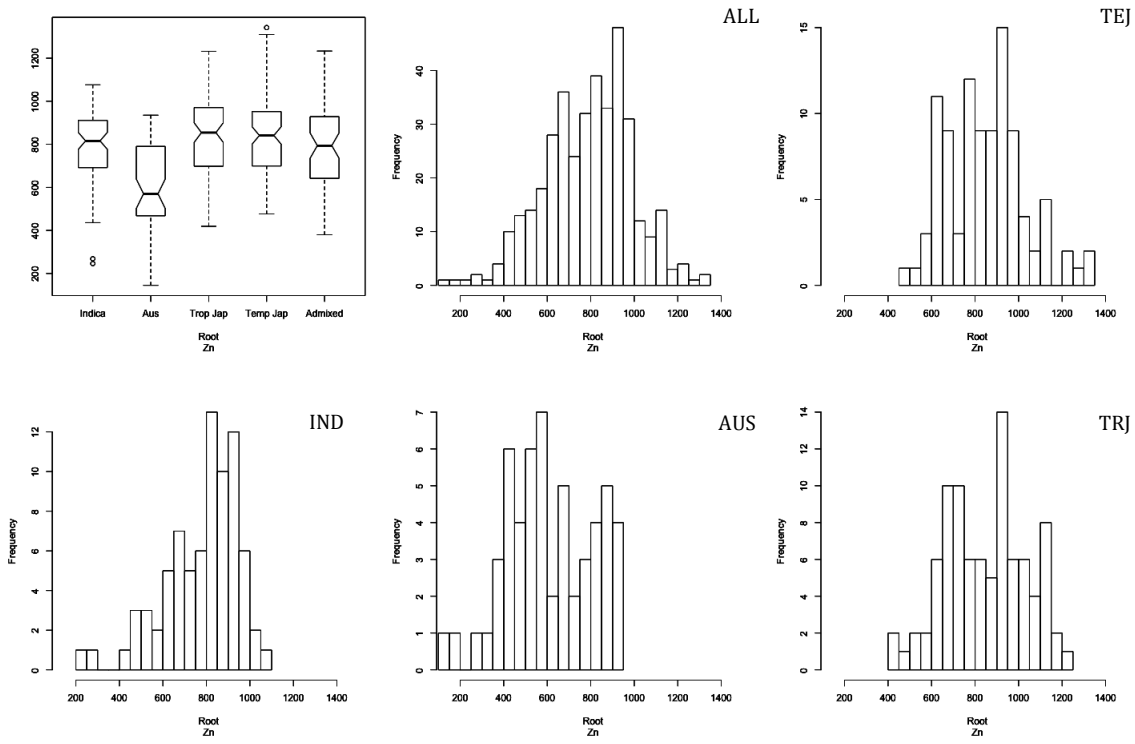
Supplementary Figure 3.96 Boxplot and histogram distributions of silicon (Si) shoot phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = 0.0027).



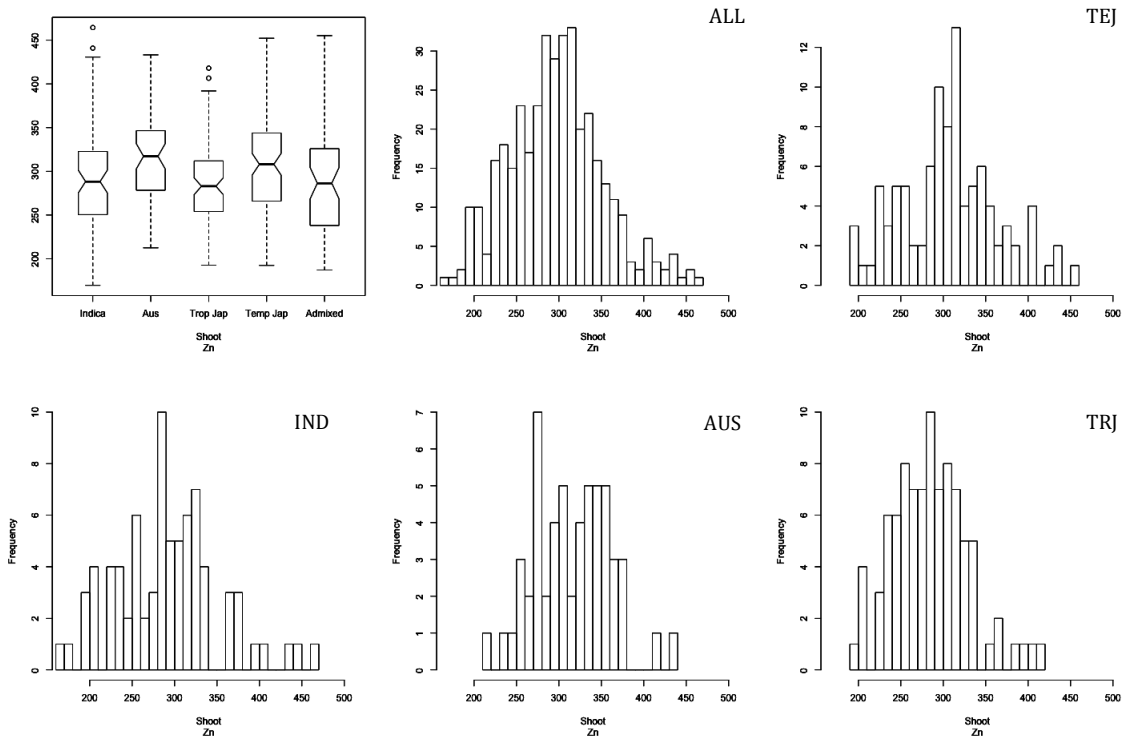
Supplementary Figure 3.97 Boxplot and histogram distributions of strontium (Sr) root phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0005).



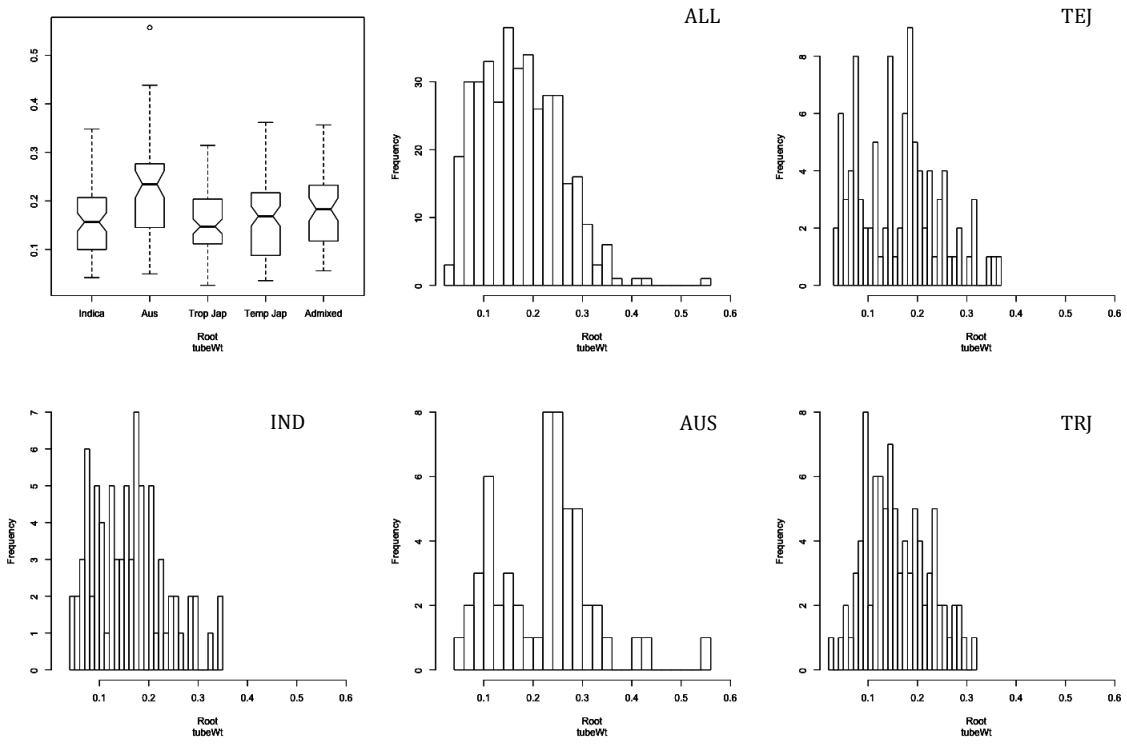
Supplementary Figure 3.98 Boxplot and histogram distributions of strontium (Sr) shoot phenotype data by subpopulation. Subpopulation structure explained 3% of the phenotypic variance (p-value = 0.0003).



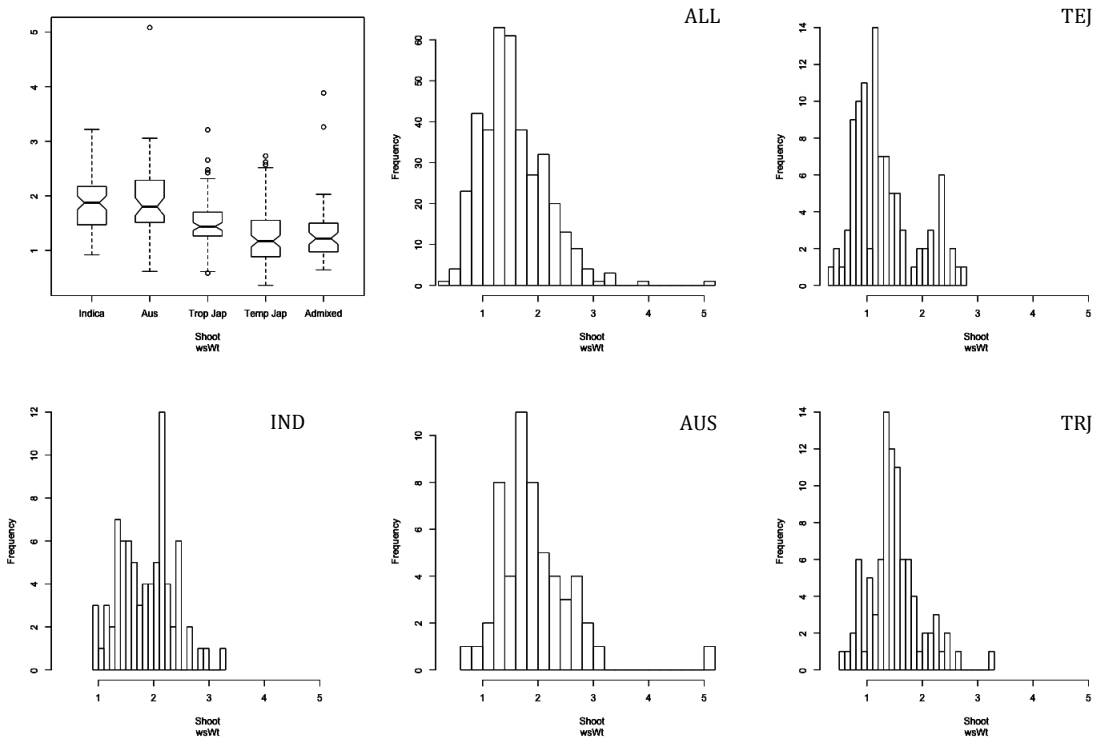
Supplementary Figure 3.99 Boxplot and histogram distributions of zinc (Zn) root phenotype data by subpopulation. Subpopulation structure explained 2% of the phenotypic variance (p-value = 0.0015).



Supplementary Figure 3.100 Boxplot and histogram distributions of zinc (Zn) shoot phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for Zn shoot content.



Supplementary Figure 3.101 Boxplot and histogram distributions of whole root dry weight biomass phenotype data by subpopulation. Subpopulation structure did not explain any significant variation for whole root dry weight biomass.



Supplementary Figure 3.102 Boxplot and histogram distributions of whole shoot dry weight biomass phenotype data by subpopulation. Subpopulation structure explained 15% of the phenotypic variance (p-value < 0.0001).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Root	1	As	TEJ	1	23315779	23643681	5.62	23366321	14
	2	B	ALL	7	22826217	22924664	4.93	22897553	6
	3	B	AUS	6	27297876	27591816	5.81	27365257	12
	4	B	IND	3	30874384	31303675	5.06	31179006	68
	5	Ca	AUS	6	665738	1290375	6.20	1125502	76
	6	Ca	IND	2	32401678	32416787	4.75	32406938	4
	7	Ca	IND	3	31126028	31303675	5.42	31298364	54
	8	Ca	IND	8	5566006	6438548	4.27	5566006	3
	9	Ca	TEJ	1	23346777	23591173	5.11	23591173	8
	10	Ca	TEJ	8	408438	1010352	7.13	885465	13
	11	Cd	TEJ	1	23338971	23643681	5.92	23523041	15
	12	Co	ALL	1	30211322	30290566	4.56	30211322	3
	13	Co	ALL	9	248437	529430	4.28	529430	3
	14	Co	ALL	10	11940350	12209074	4.52	12209074	8
	15	Cr	ALL	7	22263900	22296846	4.65	22263900	3
	16	Cr	ALL	7	22894666	22924664	4.82	22897553	4
	17	Cr	AUS	6	665738	1344591	6.47	1069445	78
	18	Cr	AUS	6	27070507	27365257	4.79	27365257	6
	19	Cr	AUS	8	636916	637101	6.15	637006	4
	20	Cr	IND	3	31179006	31298364	5.21	31265273	17
	21	Cr	TEJ	1	23338971	23624147	6.01	23591173	15
	22	Cu	AUS	2	25761069	25827209	5.28	25801640	6
	23	Cu	AUS	7	28664814	28667427	5.12	28666781	5
	24	Cu	AUS	8	25356023	25757487	5.32	25429360	11
	25	Cu	AUS	9	18783606	18925247	5.52	18857431	11
	26	Cu	AUS	11	26826261	27762679	5.32	27762679	7
	27	Cu	AUS	11	28981724	29009215	4.11	28986391	4
	28	Cu	IND	2	14148989	14817597	4.77	14148989	6
	29	Cu	TRJ	2	24378805	24459501	4.68	24394042	4
	30	Fe	ALL	7	22261483	22271232	4.94	22263900	3
	31	Fe	AUS	6	665738	1290375	6.00	1125502	53
	32	Fe	AUS	8	636916	637101	5.62	637006	3
	33	Fe	IND	3	31249805	31298364	4.25	31249805	6
	34	Fe	TEJ	1	23343063	23624147	5.60	23591173	12

Supplementary Table 3.1 Summary of GWA peaks.

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Root	35	Fe	TEJ	8	408438	1010352	5.30	885465	9
	36	I	ALL	1	29145484	29183500	4.20	29183500	3
	37	I	ALL	6	27244620	27346700	4.76	27250341	4
	38	I	AUS	6	802340	1344591	5.58	933835	41
	39	I	AUS	6	27209060	27366961	5.58	27365257	14
	40	I	AUS	9	444217	540295	4.95	444217	3
	41	I	AUS	9	1142760	1399643	4.95	1142760	3
	42	I	IND	3	31179006	31298364	4.84	31265273	13
	43	I	TEJ	8	408438	1010352	5.66	885465	5
	44	K	ALL	2	5287458	5310974	4.92	5310939	4
	45	K	ALL	6	888848	1290375	4.30	1290375	3
	46	K	ALL	6	27240812	27347992	5.29	27250341	11
	47	K	AUS	5	29607966	29628692	4.52	29628692	4
	48	K	AUS	6	933835	1344591	4.74	1271352	5
	49	K	AUS	6	27181412	27366961	5.51	27365257	40
	50	K	AUS	9	444217	540295	5.13	444217	3
	51	K	AUS	9	1142760	1399643	5.13	1142760	3
	52	K	TEJ	1	34917397	34919224	5.13	34918495	3
	53	K	TRJ	7	18658838	19992804	4.67	19991889	14
	54	Li	ALL	3	29304435	29598899	4.49	29307434	7
	55	Li	ALL	5	23162760	23242125	4.37	23162760	3
	56	Li	ALL	5	24906165	25057270	4.56	24973819	5
	57	Li	ALL	5	25874780	26980454	7.29	26250170	42
	58	Li	ALL	6	888758	1290375	5.91	1290375	12
	59	Li	AUS	6	676614	1212717	4.78	933835	18
	60	Li	AUS	6	27209060	27365257	4.85	27209060	5
	61	Li	IND	3	4732523	4751762	4.33	4732523	3
	62	Li	IND	3	31113733	31303675	5.37	31265273	22
	63	Li	TEJ	8	408438	1010352	4.52	885465	3
	64	Mg	ALL	11	7568949	7621953	5.26	7618862	5
	65	Mg	ALL	12	20046073	20149007	5.34	20058277	4
	66	Mg	ALL	12	20745551	21111225	5.39	20907529	12
	67	Mg	AUS	3	3759765	3768164	5.33	3762181	4
	68	Mg	IND	2	8290355	8375270	5.44	8293405	5
	69	Mg	IND	2	17269521	17438837	4.38	17269521	4

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Root	70	Mg	IND	5	2129008	2215618	4.39	2215618	3
	71	Mn	ALL	1	28988283	30388745	5.88	29758550	60
	72	Mn	ALL	3	28397341	29088366	6.65	29079679	7
	73	Mn	ALL	7	4853119	5135390	5.01	4853119	3
	74	Mn	ALL	7	8915536	9047298	5.62	8980684	4
	75	Mn	ALL	7	19206302	19700691	4.52	19700691	4
	76	Mn	ALL	12	24513286	24874445	7.62	24590841	7
	77	Mn	TEJ	1	23315779	23643681	8.04	23591173	21
	78	Mn	TEJ	9	5945064	6018232	5.57	5992041	4
	79	Mn	TRJ	1	29007814	29816172	5.09	29758550	6
	80	Mn	TRJ	2	28868473	29277336	4.22	28868473	3
	81	Mn	TRJ	3	28653856	29194977	5.49	29194977	6
	82	Mn	TRJ	4	34994882	35030833	4.95	34995725	3
	83	Mn	TRJ	7	4341647	6324027	8.35	5833470	23
	84	Mn	TRJ	7	6897551	9673523	12.25	8560399	29
	85	Mn	TRJ	7	10640038	13992091	5.58	11330907	20
	86	Mn	TRJ	7	14896463	16990188	8.66	16024996	17
	87	Mn	TRJ	7	19193423	19709161	6.68	19700691	27
	88	Mn	TRJ	12	24513286	25256090	7.93	24777604	31
	89	Mo	ALL	6	888848	1290375	5.78	1290375	6
	90	Mo	AUS	2	9543938	10166381	4.88	10166381	4
	91	Mo	AUS	3	18234494	19211260	4.42	19211260	3
	92	Mo	AUS	3	32980724	33077373	4.51	33077373	3
	93	Mo	AUS	4	13595581	14129605	6.14	14093560	15
	94	Mo	TEJ	1	34917357	34919224	5.05	34918495	3
	95	Mo	TEJ	11	22165958	22607665	4.94	22480416	10
	96	Na	AUS	9	3730447	4250398	4.61	3730447	3
	97	Na	IND	3	31170812	31303675	5.34	31265273	17
	98	Na	TEJ	1	23346777	23591173	4.64	23523041	4
	99	Ni	ALL	1	29145484	29183500	4.24	29183500	3
	100	Ni	ALL	6	27244620	27346700	4.78	27250341	4
	101	Ni	AUS	6	802340	1344591	5.38	1125502	38
	102	Ni	AUS	6	27209060	27366961	5.17	27365257	14
	103	Ni	AUS	9	444217	1399643	5.59	444217	7
	104	Ni	AUS	9	1904206	2500461	5.68	1904206	5

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Root	105	Ni	IND	3	31265273	31298364	4.63	31265273	5
	106	Ni	TEJ	1	23366321	23591173	4.90	23591173	4
	107	Ni	TEJ	8	371458	1010352	5.47	885465	6
	108	P	ALL	7	19554097	19557934	5.39	19556364	3
	109	P	AUS	6	802340	838315	4.49	819994	3
	110	P	IND	1	29110936	29120311	4.42	29118033	3
	111	P	IND	7	28146025	28750621	4.50	28415622	8
	112	P	TEJ	8	408438	878120	4.57	878120	3
	113	P	TEJ	9	4587405	5032349	4.40	4587405	4
	114	Pb	ALL	3	21475253	22265539	4.97	21773155	15
	115	Pb	ALL	5	3289189	3424811	4.23	3289189	3
	116	Pb	ALL	5	7873620	7942316	5.96	7880314	4
	117	Pb	ALL	11	6092665	6399212	4.62	6292903	6
	118	Pb	ALL	11	21961581	22304328	4.94	21961581	3
	119	Pb	ALL	12	26628482	26678806	4.75	26678806	5
	120	Pb	IND	1	29050028	29150350	4.92	29118033	9
	121	Pb	IND	2	6249589	6314060	4.59	6249589	5
	122	Pb	IND	5	7873620	8332608	4.66	7914363	3
	123	Pb	IND	5	25725424	25768128	4.67	25768128	9
	124	Pb	IND	7	29142588	29420306	4.58	29142588	3
	125	Pb	IND	12	9161061	9166759	4.08	9161061	4
	126	Pb	TEJ	7	1592389	1688251	5.10	1592389	3
	127	Rb	ALL	1	29122296	29499909	4.57	29183500	8
	128	Rb	ALL	6	27244620	27346700	4.77	27250341	4
	129	Rb	AUS	6	802340	1344591	5.20	933835	37
	130	Rb	AUS	6	27209060	27365593	5.10	27365257	10
	131	Rb	AUS	9	444217	540295	4.96	444217	3
	132	Rb	AUS	9	1142760	1399643	4.96	1142760	3
	133	Rb	IND	3	31179006	31298364	4.90	31265273	17
	134	Rb	TEJ	8	408438	1010352	5.54	885465	5
	135	S	ALL	3	29311021	29571958	5.27	29410804	4
	136	S	AUS	6	948665	1290375	5.50	1222551	19
	137	S	AUS	7	29409744	29414791	4.42	29414791	3
	138	S	TRJ	1	31362366	31670605	4.10	31362366	4
	139	Se	ALL	6	3148378	3243220	6.36	3148378	5

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Root	140	Se	ALL	12	24245213	24313180	4.20	24245213	9
	141	Se	AUS	1	29615659	30130607	4.83	29947864	9
	142	Se	TRJ	8	26795720	26847238	4.88	26806317	10
	143	Si	ALL	1	19438001	19687852	5.91	19438001	7
	144	Si	AUS	3	17370641	17526415	5.41	17370641	4
	145	Si	AUS	6	1048700	1069445	4.06	1058312	3
	146	Si	AUS	6	7285028	7521804	4.98	7510422	6
	147	Si	IND	1	24053209	24078116	4.30	24061337	4
	148	Si	TRJ	11	26048175	26322817	4.12	26048175	3
	149	Sr	AUS	2	25510388	25650272	5.34	25637267	4
	150	Sr	AUS	6	802340	894443	4.59	802340	8
	151	Sr	IND	1	28924884	29303536	5.25	29122296	7
	152	Sr	IND	3	31126028	31294481	5.05	31210807	37
	153	Sr	IND	6	13832318	13837968	4.19	13834831	3
	154	Sr	TEJ	1	23341359	23643681	5.83	23501367	15
	155	Sr	TEJ	8	408438	935991	4.54	885465	3
	156	Zn	ALL	5	6819739	6887945	4.96	6866993	6
	157	Zn	ALL	11	26594281	26774053	4.98	26599150	4
	158	Zn	IND	1	12280164	12422900	4.52	12298958	20
	159	Zn	IND	4	31838879	31856313	5.67	31856313	3
	160	Zn	IND	8	6708187	7320736	4.80	7161287	3
	161	Zn	TEJ	5	23563852	23627296	5.19	23609093	12
	162	Zn	TEJ	5	24346854	24583903	6.56	24471818	5
	163	Zn	TEJ	11	19233267	19425358	6.38	19425358	3

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Shoot	164	As	AUS	5	21268077	21554491	4.53	21268077	4
	165	As	IND	1	23395262	23546960	4.75	23398195	6
	166	As	IND	3	35410744	35589970	4.72	35551032	10
	167	As	TEJ	2	18353131	18587251	4.24	18390996	5
	168	As	TEJ	2	19323729	19400515	4.18	19397934	5
	169	Ca	ALL	1	10136944	10209654	4.59	10209654	3
	170	Ca	ALL	3	11275123	11376612	4.82	11376612	3
	171	Ca	IND	11	18314295	18675129	4.60	18314295	3
	172	Cd	ALL	6	29188206	29500577	5.02	29416846	4
	173	Cd	ALL	9	14814169	15013054	4.63	14814169	4
	174	Cd	AUS	1	1871	415905	4.51	292234	8
	175	Cd	AUS	2	4770994	4802039	4.33	4797798	4
	176	Cd	AUS	3	12926392	13010077	6.53	12965741	3
	177	Cd	AUS	3	35811859	35854811	5.39	35811859	5
	178	Cd	AUS	6	12974289	13339625	6.41	13202677	3
	179	Cd	AUS	7	424382	545121	6.04	495577	5
	180	Cd	AUS	7	21178877	21992234	6.54	21178877	5
	181	Cd	AUS	8	3836456	4102665	7.43	3872308	6
	182	Cd	AUS	11	20559284	20784446	7.08	20782844	3
	183	Cd	AUS	11	21972796	22474860	5.31	22474860	7
	184	Cd	TRJ	1	861730	1032090	5.07	861730	5
	185	Cd	TRJ	6	29588432	29642473	4.58	29588432	3
	186	Co	AUS	4	18936104	19208524	4.97	18936104	4
	187	Co	TEJ	3	14810243	15583631	4.57	14832862	14
	188	Co	TEJ	4	13682646	15084915	4.23	14214329	9
	189	Co	TRJ	7	28706880	29248197	4.45	29248197	4
	190	Cr	ALL	3	12931369	13007155	4.98	12970471	4
	191	Cr	IND	6	1550846	2126533	4.79	1843333	19
	192	Cr	IND	9	19274086	19610167	4.87	19610167	8
	193	Cu	ALL	3	5777754	5796251	4.90	5777754	3
	194	Cu	ALL	4	4869256	5105482	4.72	4936961	5
	195	Cu	AUS	9	15665247	15880023	4.59	15880023	3
	196	Cu	IND	4	17583307	18570914	4.64	17583307	5
	197	Cu	IND	8	4671983	5038925	5.30	4767925	21
	198	Cu	IND	10	208457	321214	4.53	321214	3
	199	Cu	IND	11	22153562	22433640	5.15	22433640	13

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Shoot	200	Fe	ALL	6	27244620	27365257	6.59	27250341	7
	201	Fe	AUS	6	27245820	27365257	4.93	27250341	4
	202	Fe	IND	3	35912262	36243919	5.05	36241743	4
	203	Fe	IND	6	26886515	27106735	4.79	26886515	5
	204	Fe	TEJ	7	22510121	23178153	4.95	22515424	7
	205	I	ALL	7	18508120	18654861	5.40	18654861	4
	206	I	AUS	8	16922	85414	4.36	16922	9
	207	I	IND	4	14209773	15080495	5.33	14540875	10
	208	I	IND	4	16296552	16494299	4.49	16494299	7
	209	I	IND	5	3890377	4125320	4.20	4125320	6
	210	I	TEJ	1	22980639	23084851	4.37	22987492	3
	211	I	TEJ	1	32103912	32260885	4.37	32111639	4
	212	I	TEJ	7	22515424	22924664	4.77	22916565	5
	213	K	ALL	1	38242293	38335941	4.45	38294582	4
	214	K	AUS	1	4622619	5229097	4.57	4622619	4
	215	K	AUS	1	27995687	30115859	4.97	28710381	8
	216	K	AUS	1	39507550	40119743	5.05	39940070	4
	217	K	AUS	2	3090966	4374988	4.94	4374988	4
	218	K	AUS	2	7599316	8292903	5.39	7599316	3
	219	K	AUS	2	10664141	11218910	4.67	11218910	4
	220	K	AUS	2	14737312	15330207	4.98	15330207	3
	221	K	AUS	2	26737324	27600091	5.01	26737324	5
	222	K	AUS	2	31889061	32720360	5.22	32024610	6
	223	K	AUS	2	35525737	35772744	5.18	35525737	3
	224	K	AUS	3	7505803	8463355	5.27	7505803	5
	225	K	AUS	3	9021632	9564556	5.07	9021632	3
	226	K	AUS	3	19181935	20111237	5.01	19868909	5
	227	K	AUS	3	30916945	31070960	4.96	30916945	3
	228	K	AUS	3	32188258	32634755	5.07	32625274	3
	229	K	AUS	4	154487	590469	5.43	278717	6
	230	K	AUS	4	24184803	24477797	5.35	24278353	3
	231	K	AUS	4	28467312	29086219	4.28	28467312	3
	232	K	AUS	4	30122230	31330466	5.47	31330466	5
	233	K	AUS	4	35131240	35216258	5.27	35141825	3
	234	K	AUS	5	3503460	3786191	4.98	3503460	3

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Shoot	235	K	AUS	5	11094735	11504562	5.51	11094735	3
	236	K	AUS	5	13007423	13810556	4.83	13314069	4
	237	K	AUS	5	18069024	18951827	4.95	18479954	5
	238	K	AUS	5	22999721	23653123	4.47	23653123	3
	239	K	AUS	5	25146331	25700676	4.99	25700676	4
	240	K	AUS	5	27743719	27949052	4.84	27933681	3
	241	K	AUS	6	6329293	6573980	5.40	6329293	4
	242	K	AUS	6	7562663	8017951	4.78	7562663	4
	243	K	AUS	6	12784098	13475269	4.59	13099667	4
	244	K	AUS	6	16617638	17322917	5.17	16864925	4
	245	K	AUS	6	22993345	23328083	5.01	23167924	4
	246	K	AUS	6	26738864	26883500	4.21	26738864	3
	247	K	AUS	7	1873795	2496746	5.03	2050379	3
	248	K	AUS	7	3230562	3406619	4.93	3406619	3
	249	K	AUS	7	6054211	6223152	5.08	6161839	3
	250	K	AUS	8	3246377	4068565	4.53	3611473	3
	251	K	AUS	8	10730232	10955673	4.96	10730232	3
	252	K	AUS	9	9605468	9993344	5.09	9605468	3
	253	K	AUS	9	17114488	18143973	5.38	17114488	4
	254	K	AUS	10	4510121	4989594	4.57	4510121	3
	255	K	AUS	10	15139740	15829889	5.08	15623501	3
	256	K	AUS	10	17070954	17847334	4.95	17070954	5
	257	K	AUS	10	18440701	19592406	5.14	19592406	7
	258	K	AUS	11	2825707	3670795	4.98	2825707	5
	259	K	AUS	11	6543099	6834229	4.96	6543099	7
	260	K	AUS	11	20379701	21411524	5.40	20481479	5
	261	K	AUS	11	24940262	25174859	5.02	24965186	3
	262	K	AUS	12	19397447	19801415	5.42	19801415	3
	263	K	AUS	12	23479599	23884421	4.86	23884421	3
	264	K	AUS	12	26023144	26614915	4.94	26346396	3
	265	K	IND	1	29463693	30339697	6.77	29463693	7
	266	K	IND	1	32950985	34129689	5.66	33319349	4
	267	K	IND	1	42588487	42633864	5.60	42620042	4
	268	K	IND	2	21086603	21590332	4.37	21590332	4
	269	K	IND	2	26069482	26682220	4.66	26069482	3

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Shoot	270	K	IND	4	58864	458929	4.08	58864	3
	271	K	IND	6	7692740	7760675	4.49	7760675	3
	272	K	IND	6	8894604	9477259	4.94	8894604	7
	273	K	IND	8	527192	560290	4.13	555392	6
	274	K	IND	8	27425634	27551266	5.47	27474258	3
	275	K	IND	11	21172654	21323271	6.68	21323271	3
	276	K	IND	12	8765565	9292944	4.28	9193540	3
	277	K	IND	12	11588371	12341164	6.13	11703999	5
	278	Li	ALL	6	22678090	23214440	5.13	23214440	4
	279	Li	ALL	7	11475186	11704702	4.51	11475186	3
	280	Li	ALL	7	12413100	13281653	4.30	13281653	4
	281	Li	IND	2	4561293	4572528	4.33	4561293	3
	282	Li	IND	10	6273257	6695706	4.42	6273257	3
	283	Li	IND	10	9172744	9794490	4.43	9794490	4
	284	Mg	ALL	11	17086644	17379890	4.19	17316500	3
	285	Mg	AUS	1	35343330	35494441	4.40	35343330	3
	286	Mg	IND	1	22640125	22913650	4.99	22640125	8
	287	Mg	IND	3	2273913	2298354	4.69	2273913	3
	288	Mg	IND	8	8772106	8805623	5.04	8772106	7
	289	Mg	TEJ	6	3562814	3602700	5.25	3602700	9
	290	Mn	ALL	1	34817951	35093428	5.68	34892065	9
	291	Mn	ALL	10	1925788	2019805	6.85	1965441	10
	292	Mn	AUS	4	5316204	5320155	4.22	5316204	3
	293	Mn	AUS	10	1632107	2086124	5.66	1812624	5
	294	Mn	IND	4	18041132	18950984	5.12	18950984	10
	295	Mn	TRJ	11	5370874	5709028	5.54	5370874	3
	296	Mo	ALL	8	15243	648977	15.39	200233	152
	297	Mo	AUS	10	17225629	17468081	4.82	17225629	3
	298	Mo	IND	1	9769265	10029348	4.05	10029348	4
	299	Mo	IND	1	12868405	13435952	4.58	13228832	6
	300	Mo	IND	8	9952	118947	4.64	116154	10
	301	Mo	TRJ	2	18623912	18630619	5.70	18626031	6
	302	Mo	TRJ	8	16922	320915	6.37	200233	69
	303	Na	ALL	8	7451726	7518485	4.93	7488138	12
	304	Na	ALL	8	27470977	28060458	7.55	27504822	27

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Shoot	305	Na	ALL	9	21706927	21961224	5.49	21926446	7
	306	Na	AUS	1	2173566	3784649	8.43	2173566	18
	307	Na	AUS	1	20436138	20996673	4.70	20436138	3
	308	Na	AUS	1	21625580	21660511	4.39	21625580	4
	309	Na	AUS	1	33427940	33783082	4.74	33611031	3
	310	Na	AUS	4	5446315	5784232	6.72	5784232	3
	311	Na	AUS	4	30548798	31148129	8.46	31027452	63
	312	Na	AUS	8	6340786	7463918	5.76	7132839	7
	313	Na	AUS	8	20552012	21076009	4.53	20560046	6
	314	Na	AUS	8	27468044	28059167	8.41	27585649	28
	315	Na	AUS	11	1557071	1571203	4.62	1571203	6
	316	Na	AUS	11	25787205	25829801	4.42	25814569	5
	317	Na	IND	2	24840284	25009820	4.93	24922063	5
	318	Na	IND	5	7759436	7778665	5.74	7770039	3
	319	Na	IND	6	21162563	21198668	4.26	21165004	5
	320	Na	IND	6	21999646	22013011	5.48	22002438	4
	321	Na	IND	6	23726184	23908959	6.25	23880293	10
	322	Na	TEJ	8	7737265	8305931	4.23	7941068	3
	323	Ni	ALL	4	33732514	33835671	4.50	33780362	8
	324	Ni	ALL	5	24973819	25206799	6.18	25054618	10
	325	Ni	ALL	7	19958320	20561696	5.59	20084671	8
	326	Ni	AUS	1	18690404	18725405	4.25	18707335	3
	327	Ni	AUS	6	7313739	7478120	5.15	7414753	11
	328	Ni	IND	7	22680383	23231066	5.96	23229073	4
	329	P	ALL	1	2001792	2002635	5.12	2002080	3
	330	P	TEJ	1	1874689	2061535	6.63	2001792	79
	331	P	TEJ	2	9661058	9776692	4.38	9716094	4
	332	Pb	ALL	6	25887619	26063255	4.39	25887619	3
	333	Pb	ALL	7	9115025	9313052	5.01	9196108	3
	334	Pb	TEJ	1	23338971	23523041	4.65	23343063	9
	335	Rb	AUS	2	20422375	21052036	4.22	21022136	12
	336	Rb	TEJ	6	10494286	10545225	5.28	10532008	3
	337	S	ALL	1	25269849	25293079	5.26	25277070	7
	338	S	ALL	3	4954427	5482615	4.52	5468596	5
	339	S	ALL	5	4984765	5174807	8.34	5110523	8

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Ion	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position	# of significant SNPs in region
Shoot	340	S	ALL	5	16376289	16481639	4.66	16481639	3
	341	S	ALL	5	19956751	20436357	4.19	20436357	3
	342	S	ALL	9	5971601	6598973	4.94	6261692	3
	343	S	AUS	3	1833744	1839135	4.87	1834615	3
	344	S	TEJ	1	24607600	24774092	4.08	24723609	5
	345	S	TEJ	5	20434408	20439744	5.66	20434408	4
	346	S	TEJ	9	9077744	9203668	4.34	9077744	7
	347	S	TRJ	5	13828640	13945193	4.27	13828640	3
	348	Se	ALL	1	41244480	41587329	4.72	41320564	8
	349	Se	ALL	12	17654666	18121845	5.28	18121845	4
	350	Se	IND	6	725500	744876	5.66	742145	6
	351	Se	IND	7	7002502	7239622	5.56	7194338	27
	352	Se	TRJ	12	17015815	17395209	4.41	17015815	5
	353	Si	ALL	10	1711242	1884948	4.75	1711242	3
	354	Si	AUS	4	5292411	5320155	4.70	5316720	11
	355	Sr	ALL	4	29686878	29749647	4.57	29686878	3
	356	Sr	ALL	5	25839893	25939552	4.87	25839893	3
	357	Sr	IND	3	23442785	23472227	4.12	23471146	4
	358	Zn	AUS	3	13170835	13707521	4.09	13170835	4
	359	Zn	AUS	4	22713197	22793211	4.18	22713197	3
	360	Zn	AUS	5	7397430	7468733	4.50	7468733	4
	361	Zn	AUS	8	19282078	19575921	5.19	19575921	19
	362	Zn	TRJ	6	14407847	15148145	4.59	15148145	3

Supplementary Table 3.1 (continued).

Tissue	Peak ID	Phenotype	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position (bp)	Number of significant SNPs in region
Root	1	DW Biomass	ALL	11	22175445	22229352	5.31	22218247	9
	2	DW Biomass	ALL	12	16572914	16573422	4.38	16572914	3
	3	DW Biomass	AUS	11	21300178	21958520	4.73	21300178	4
	4	DW Biomass	IND	4	14188745	14210684	5.14	14188812	6
	5	DW Biomass	TEJ	5	8019201	8125496	5.91	8125496	4
	6	DW Biomass	TEJ	7	22526139	23104027	4.89	22526139	9
	7	DW Biomass	TEJ	12	16452234	16579036	5.30	16505204	12
Shoot	8	DW Biomass	ALL	1	23110137	24172561	5.28	23636753	5
	9	DW Biomass	ALL	3	28662027	28910693	5.36	28689536	5
	10	DW Biomass	ALL	4	11280520	11406913	4.54	11406913	3
	11	DW Biomass	ALL	6	27346190	27403764	5.40	27346190	8
	12	DW Biomass	ALL	7	28290387	28924479	6.29	28292381	7
	13	DW Biomass	ALL	8	20583349	20656448	4.62	20583349	5
	14	DW Biomass	ALL	8	27431783	28390513	7.09	28390513	24
	15	DW Biomass	AUS	1	7000762	7844665	5.89	7000762	4
	16	DW Biomass	AUS	1	12053191	13297675	6.00	12080365	6
	17	DW Biomass	AUS	1	21947697	22627369	5.56	22448851	5
	18	DW Biomass	AUS	1	30575005	31569396	4.78	31569396	5
	19	DW Biomass	AUS	2	19772390	20027636	5.47	19772390	3
	20	DW Biomass	AUS	2	23827320	24342407	6.29	24244675	3
	21	DW Biomass	AUS	3	14026004	14472863	4.79	14472863	3
	22	DW Biomass	AUS	3	23996007	24469905	5.96	24200942	6
	23	DW Biomass	AUS	4	24524127	25113370	5.89	24524127	5
	24	DW Biomass	AUS	5	13447503	13856067	4.60	13729011	3
	25	DW Biomass	AUS	5	20853698	21391523	4.47	20910584	3
	26	DW Biomass	AUS	5	25583648	26395069	4.86	25583648	3
	27	DW Biomass	AUS	6	4733353	5460719	4.41	5460719	3
	28	DW Biomass	AUS	6	16541699	17321328	5.59	16541699	3
	29	DW Biomass	AUS	7	212171	722663	6.30	722663	3
	30	DW Biomass	AUS	8	19444792	20061445	6.23	19793379	3
	31	DW Biomass	AUS	8	26456867	27894292	5.38	26456867	4
	32	DW Biomass	AUS	9	12545322	13157056	5.58	12545322	3
	33	DW Biomass	AUS	10	5554612	6543196	5.17	5554612	5
	34	DW Biomass	AUS	10	13058762	13436291	6.12	13058762	3

Supplementary Table 3.2 Summary of GWA peaks for root and shoot biomass.

Tissue	Peak ID	Phenotype	Subpop	Chr	Start (bp) MSU7	Stop (bp) MSU7	Max Negative LOG(p-value)	Peak SNP Position (bp)	Number of significant SNPs in region
Shoot	35	DW Biomass	AUS	10	17000410	17786196	5.54	17651536	4
	36	DW Biomass	AUS	11	5246181	5317331	6.18	5317331	3
	37	DW Biomass	AUS	11	11048660	11252684	5.10	11048660	3
	38	DW Biomass	AUS	11	18984703	19862475	5.82	19862475	4
	39	DW Biomass	AUS	12	9427757	10082280	4.96	9427757	3
	40	DW Biomass	AUS	12	11752377	12700420	5.63	11752377	4
	41	DW Biomass	AUS	12	14604896	15094700	5.02	15094700	3
	42	DW Biomass	AUS	12	18305855	18837378	5.61	18837378	3
	43	DW Biomass	IND	3	35695364	35747828	4.84	35743038	5
	44	DW Biomass	IND	6	6996839	7552457	4.52	7549158	7
	45	DW Biomass	IND	12	14520563	14989923	4.66	14985444	7
	46	DW Biomass	TEJ	5	18010808	18109068	5.55	18018686	14
	47	DW Biomass	TEJ	7	6342461	6378388	4.22	6342461	5
	48	DW Biomass	TEJ	8	8119447	8677013	6.10	8667086	6
	49	DW Biomass	TEJ	10	3013496	3384210	5.70	3017423	6
	50	DW Biomass	TEJ	11	25524147	26090875	5.67	25596026	4
	51	DW Biomass	TRJ	3	10160492	10186052	6.56	10185539	3

Supplementary Table 3.2 (continued).

Locus ID	Chr	Ion	Start (bp) MSU7	Stop (bp) MSU7	Hit ID from Table S3.1	Tissue	Subpop
LOC_Os06g01170	6	Ca	130089	137360	26	root	AUS
LOC_Os06g01690	6	Ca	409926	412145	26	root	AUS
LOC_Os06g01760	6	Ca	437685	440554	26	root	AUS
LOC_Os06g01890	6	Ca	493851	510407	26	root	AUS
LOC_Os03g54100	3	Ca	31007714	31010336	28	root	IND
LOC_Os03g54160	3	Ca	31024033	31034456	28	root	IND
LOC_Os03g54170	3	Ca	31041238	31047906	28	root	IND
LOC_Os03g55100	3	Ca	31337819	31343721	28	root	IND
LOC_Os08g02070	8	Ca	678358	680739	29	root	IND
LOC_Os08g02070	8	Ca	678358	680739	31	root	TEJ
LOC_Os01g18890	1	Ca	10678464	10681164	35	shoot	ALL
LOC_Os06g48720	6	Cd	29476951	29479907	40	shoot	ALL
LOC_Os02g08010	2	Cd	4195421	4202350	43	shoot	AUS
LOC_Os02g08014	2	Cd	4202296	4202592	43	shoot	AUS
LOC_Os02g08018	2	Cd	4203438	4208367	43	shoot	AUS
LOC_Os08g06460	8	Cd	3621707	3624002	49	shoot	AUS
LOC_Os06g48720	6	Cd	29476951	29479907	53	shoot	TRJ
LOC_Os04g32030	4	Co	19017158	19020428	57	shoot	AUS
LOC_Os03g27040	3	Co	15478390	15480786	58	shoot	TEJ
LOC_Os04g25470	4	Co	14752994	14760563	59	shoot	TEJ
YSL2	2	Cu	26164518	26169101	72	root	AUS
LOC_Os07g47480	7	Cu	28381565	28383153	73	root	AUS
LOC_Os11g47809	11	Cu	28304455	28305292	77	root	AUS
LOC_Os02g39230	2	Cu	23694091	23695761	79	root	TRJ
LOC_Os02g39240	2	Cu	23699789	23700991	79	root	TRJ
LOC_Os02g39260	2	Cu	23705021	23706577	79	root	TRJ
LOC_Os02g39420	2	Cu	23778892	23780064	79	root	TRJ
LOC_Os02g39430	2	Cu	23783066	23784268	79	root	TRJ
LOC_Os03g11030	3	Cu	5672158	5675258	81	shoot	ALL
LOC_Os03g11040	3	Cu	5676634	5679112	81	shoot	ALL
LOC_Os09g25314	9	Cu	15155994	15158085	83	shoot	AUS
LOC_Os06g03770	6	Fe	1491884	1501598	89	root	AUS
LOC_Os06g04020	6	Fe	1641215	1642801	89	root	AUS
LOC_Os03g54910	3	Fe	31209286	31211707	91	root	IND

Supplementary Table 3.3 Cloned and putative candidate genes that aligned with GWA peaks.

Locus ID	Chr	Ion	Start (bp) MSU7	Stop (bp) MSU7	Hit ID from Table S3.1	Tissue	Subpop
LOC_Os06g45450	6	Fe	27472218	27473452	97	shoot	ALL
LOC_Os06g46310	6	Fe	28058356	28064182	97	shoot	ALL
LOC_Os06g45450	6	Fe	27472218	27473452	98	shoot	AUS
LOC_Os06g46310	6	Fe	28058356	28064182	98	shoot	AUS
LOC_Os03g64140	3	Fe	36235979	36238006	99	shoot	IND
LOC_Os03g64340	3	Fe	36359238	36360640	99	shoot	IND
LOC_Os06g45450	6	Fe	27472218	27473452	100	shoot	IND
LOC_Os07g38910	7	Fe	23329028	23329853	101	shoot	TEJ
LOC_Os02g09480	2	K	4867671	4869377	119	root	ALL
LOC_Os06g03610	6	K	1396922	1399969	120	root	ALL
LOC_Os06g44280	6	K	26713869	26716573	121	root	ALL
LOC_Os06g44460	6	K	26853888	26857589	121	root	ALL
LOC_Os06g45940	6	K	27818966	27826495	121	root	ALL
LOC_Os05g51090	5	K	29250342	29252223	122	root	AUS
LOC_Os05g51610	5	K	29526001	29531196	122	root	AUS
LOC_Os06g03610	6	K	1396922	1399969	123	root	AUS
LOC_Os06g44280	6	K	26713869	26716573	124	root	AUS
LOC_Os06g44460	6	K	26853888	26857589	124	root	AUS
LOC_Os06g45940	6	K	27818966	27826495	124	root	AUS
LOC_Os09g02020	9	K	703295	709944	125	root	AUS
LOC_Os09g02214	9	K	877145	889186	126	root	AUS
LOC_Os09g02020	9	K	703295	709944	126	root	AUS
LOC_Os09g02214	9	K	877145	889186	126	root	AUS
LOC_Os01g59490	1	K	34405333	34408175	127	root	TEJ
LOC_Os01g60140	1	K	34787358	34790577	127	root	TEJ
LOC_Os07g30960	7	K	18330762	18332965	128	root	TRJ
LOC_Os07g32530	7	K	19377446	19385145	128	root	TRJ
LOC_Os07g32680	7	K	19513897	19515320	128	root	TRJ
LOC_Os07g32700	7	K	19528479	19529369	128	root	TRJ
LOC_Os07g32710	7	K	19532183	19533584	128	root	TRJ
LOC_Os07g33350	7	K	19934500	19935430	128	root	TRJ
LOC_Os07g33790	7	K	20220084	20224931	128	root	TRJ
LOC_Os07g39310	7	K	23539714	23543082	128	root	TRJ
LOC_Os01g66170	1	K	38435799	38439417	129	shoot	ALL
LOC_Os01g66420	1	K	38565104	38569818	129	shoot	ALL

Supplementary Table 3.3 (continued).

Locus ID	Chr	Ion	Start (bp) MSU7	Stop (bp) MSU7	Hit ID from Table S3.1	Tissue	Subpop
LOC_Os01g51420	1	K	29567502	29570439	181	shoot	IND
LOC_Os01g52070	1	K	29942865	29947609	181	shoot	IND
LOC_Os01g53294	1	K	30621627	30633361	181	shoot	IND
LOC_Os01g56330	1	K	32460834	32464022	182	shoot	IND
LOC_Os01g57890	1	K	33470856	33477491	182	shoot	IND
LOC_Os01g58380	1	K	33729334	33731992	182	shoot	IND
LOC_Os01g58740	1	K	33951049	33955085	182	shoot	IND
LOC_Os01g59490	1	K	34405333	34408175	182	shoot	IND
LOC_Os01g73810	1	K	42756787	42757563	183	shoot	IND
LOC_Os01g74470	1	K	43124919	43133458	183	shoot	IND
LOC_Os02g35300	2	K	21219175	21222864	184	shoot	IND
LOC_Os02g35320	2	K	21237383	21240893	184	shoot	IND
LOC_Os02g35338	2	K	21247879	21249767	184	shoot	IND
LOC_Os02g35356	2	K	21256905	21258793	184	shoot	IND
LOC_Os02g35374	2	K	21265931	21267819	184	shoot	IND
LOC_Os02g35392	2	K	21274957	21276845	184	shoot	IND
LOC_Os02g35420	2	K	21283983	21285873	184	shoot	IND
LOC_Os02g35600	2	K	21408016	21411353	184	shoot	IND
LOC_Os02g36330	2	K	21921698	21923050	184	shoot	IND
LOC_Os02g43300	2	K	26093366	26097273	185	shoot	IND
LOC_Os02g43540	2	K	26265810	26266788	185	shoot	IND
LOC_Os04g01600	4	K	399955	404335	186	shoot	IND
LOC_Os04g01650	4	K	442307	443840	186	shoot	IND
LOC_Os04g01660	4	K	445010	446108	186	shoot	IND
LOC_Os04g01674	4	K	452417	456130	186	shoot	IND
LOC_Os06g14030	6	K	7798074	7805572	187	shoot	IND
LOC_Os06g14240	6	K	7939907	7940693	187	shoot	IND
LOC_Os06g14310	6	K	7981085	7982860	187	shoot	IND
LOC_Os06g15150	6	K	8583739	8589596	188	shoot	IND
LOC_Os06g15620	6	K	8846703	8847848	188	shoot	IND
LOC_Os06g15910	6	K	9032863	9036776	188	shoot	IND
LOC_Os08g02050	8	K	661168	666559	189	shoot	IND
LOC_Os08g02450	8	K	980212	983446	189	shoot	IND
LOC_Os08g42920	8	K	27125046	27127010	190	shoot	IND
LOC_Os08g43690	8	K	27616156	27618880	190	shoot	IND

Supplementary Table 3.3 (continued).

Locus ID	Chr	Ion	Start (bp) MSU7	Stop (bp) MSU7	Hit ID from Table S3.1	Tissue	Subpop
LOC_Os08g44050	8	K	27730405	27736989	190	shoot	IND
LOC_Os08g44340	8	K	27897097	27900409	190	shoot	IND
LOC_Os12g16250	12	K	9288203	9288904	192	shoot	IND
LOC_Os12g16290	12	K	9312637	9315532	192	shoot	IND
LOC_Os03g04480	3	Mg	2074318	2076632	220	shoot	IND
LOC_Os03g05750	3	Mg	2865311	2867238	220	shoot	IND
LOC_Os01g53210	1	Mn	30573817	30575243	223	root	ALL
LOC_Os07g15370	7	Mn	8870442	8877911	226	root	ALL
LOC_Os07g15460	7	Mn	8965029	8969888	226	root	ALL
LOC_Os12g39180	12	Mn	24086315	24090178	228	root	ALL
ZIP8	7	Mn	7393491	7396720	236	root	TRJ
LOC_Os07g15370	7	Mn	8870442	8877911	236	root	TRJ
LOC_Os07g15460	7	Mn	8965029	8969888	236	root	TRJ
LOC_Os08g01120	8	Mo	85335	87510	254	shoot	ALL
LOC_Os08g01120	8	Mo	85335	87510	258	shoot	IND
LOC_Os08g01120	8	Mo	85335	87510	260	shoot	TRJ
LOC_Os09g07500	9	Na	3750007	3753991	261	root	AUS
LOC_Os03g53630	3	Na	30740482	30748610	262	root	IND
LOC_Os03g54160	3	Na	31024033	31034456	262	root	IND
LOC_Os03g54170	3	Na	31041238	31047906	262	root	IND
LOC_Os01g37690	1	Na	21070325	21075685	263	root	TEJ
LOC_Os08g44310	8	Na	27885222	27886124	265	shoot	ALL
LOC_Os08g43690	8	Na	27616156	27618880	265	shoot	ALL
LOC_Os01g05790	1	Na	2764773	2767644	267	shoot	AUS
LOC_Os01g37690	1	Na	21070325	21075685	268	shoot	AUS
LOC_Os01g34850	1	Na	19242042	19243853	268	shoot	AUS
LOC_Os04g51610	4	Na	30386970	30399183	272	shoot	AUS
OsHKT1;2	4	Na	30548314	30545885	272	shoot	AUS
LOC_Os04g51820	4	Na	30724244	30727084	272	shoot	AUS
LOC_Os04g51830	4	Na	30734183	30739334	272	shoot	AUS
LOC_Os08g34120	8	Na	21390687	21392300	274	shoot	AUS
LOC_Os08g44310	8	Na	27885222	27886124	275	shoot	AUS
LOC_Os08g43690	8	Na	27616156	27618880	275	shoot	AUS
LOC_Os11g04460	11	Na	1858334	1864303	276	shoot	AUS
LOC_Os06g36940	6	Na	21762121	21764047	280	shoot	IND

Supplementary Table 3.3 (continued).

Locus ID	Chr	Ion	Start (bp) MSU7	Stop (bp) MSU7	Hit ID from Table S3.1	Tissue	Subpop
LOC_Os06g36940	6	Na	21762121	21764047	281	shoot	IND
LOC_Os06g45500	6	Ni	27516103	27522606	285	root	ALL
LOC_Os06g02580	6	Ni	907486	909323	286	root	AUS
LOC_Os06g45500	6	Ni	27516103	27522606	287	root	AUS
LOC_Os09g03310	9	Ni	1611177	1614358	288	root	AUS
LOC_Os09g03310	9	Ni	1611177	1614358	289	root	AUS
LOC_Os03g41064	3	Pb	22822261	22823032	308	root	ALL
LOC_Os03g41070	3	Pb	22824386	22825030	308	root	ALL
ZIP6	5	Pb	3807951	3810783	309	root	ALL
LOC_Os07g15370	7	Pb	8870442	8877911	322	shoot	ALL
LOC_Os07g15460	7	Pb	8965029	8969888	322	shoot	ALL
LOC_Os03g09930	3	S	4945325	4952952	346	shoot	ALL
LOC_Os03g09940	3	S	4953103	4958450	346	shoot	ALL
LOC_Os03g09970	3	S	4983579	4991413	346	shoot	ALL
LOC_Os03g09980	3	S	4995775	5001179	346	shoot	ALL
LOC_Os03g10500	3	S	5340101	5346082	346	shoot	ALL
LOC_Os11g43860	11	Zn	26016786	26020797	436	root	ALL
ZIP4	8	Zn	6266828	6269910	439	root	IND
LOC_Os05g39540	5	Zn	23141972	23150720	440	root	TEJ
ZIP5	5	Zn	23153854	23156724	440	root	TEJ

Supplementary Table 3.3 (continued).

Admixed Line	*O. rufipogon* genomic introgressions [Position in MB (size in MB)]			
	Chromosome 1	Chromosome 2	Chromosome 3	Chromosome 4
141_2-1	0.6-20.2 MB (19.4)	9.5-36.6 MB (27.1)	33.9-36.3 MB (2.4)	*none*
141_2-6	7.8-20.2 MB (12.4)	5.7-16.7 MB (11)	*none*	*none*
141_2-11	*none*	5.7-16.7 MB (11)	*none*	30.7 - 32.4 MB (1.7)

	Chromosome 5	Chromosome 9	Chromosome 11	Chromosome 12
141_2-1	28.9-29.7 MB (0.8)	17.8-23.2 MB (5.4)	17.9-30.6 MB (12.7)	14.2-24.8 MB (10.6)
141_2-6	27.9-29.7 MB (1.8)	17.8-23.2 MB (5.4)	*none*	*none*
141_2-11	27.9-29.7 MB (1.8)	22.0-23.2 MB (1.2)	*none*	*none*

Supplementary Table 3.4 Summary of *O. rufipogon* background introgressions in near-isogenic lines.

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
1	168.189	-14.540	5663.487	118.428	38.323	23.124	145.956	14527.702
2	159.918	-1.078	5626.124	107.585	31.616	21.746	91.371	13609.265
3	193.208	-5.837	5113.259	121.273	34.116	22.712	196.794	13362.252
4	135.734	-7.914	5148.670	125.927	35.287	21.327	84.375	12899.524
5	209.421	-19.000	5330.434	108.012	34.742	27.201	170.426	14535.685
6	160.624	2.085	5550.922	144.968	38.270	22.331	111.701	15132.296
7	188.937	35.529	5750.993	127.271	39.700	24.064	221.864	14223.572
8	195.225	-18.216	6272.783	127.897	40.558	27.001	97.811	16637.575
9	216.584	-14.568	5963.597	138.561	52.573	25.896	267.605	17143.291
10	188.288	4.664	5531.474	148.062	46.058	23.986	189.477	17405.081
11	194.044	9.676	5755.384	156.329	37.779	25.993	250.283	15930.728
12	162.410	-2.102	5460.337	101.639	38.912	23.042	166.015	13452.206
13	189.681	-7.197	5632.644	88.375	46.748	22.984	91.501	15429.506
14	246.649	-5.821	6060.201	149.036	34.577	37.123	312.896	18779.906
15	185.482	21.709	5313.716	140.423	48.379	27.195	185.414	15082.936
16	186.282	-22.161	5851.136	104.072	45.765	27.568	171.741	15318.567
17	232.595	34.559	5920.318	159.403	42.824	31.395	267.428	17626.531
18	267.102	4.416	6568.408	199.326	60.395	27.629	207.834	19914.854
19	161.558	-21.550	5156.734	94.638	36.315	20.417	107.634	13821.866
20	167.794	-2.097	5570.415	142.586	45.038	25.205	170.652	14396.974
21	207.150	31.387	5065.732	168.418	52.283	23.805	213.909	15004.784
22	142.062	-23.492	5270.431	107.510	40.446	21.004	142.341	13806.010
23	169.221	-15.102	5997.503	146.830	47.449	25.523	167.051	14227.468
24	202.180	-7.294	5992.760	135.920	46.653	24.424	139.437	14023.254
25	225.674	-6.395	6473.446	132.350	44.978	28.806	138.000	16066.908
26	233.720	-16.298	6033.567	127.213	46.066	25.319	115.338	14947.714
27	163.852	-10.389	6257.240	153.999	46.660	25.801	128.493	15058.280
28	151.590	-27.910	5192.788	125.077	33.781	20.937	137.616	12342.105
29	217.916	3.340	6259.292	146.281	38.691	26.585	197.346	15699.835
30	213.476	8.635	5955.501	189.493	46.343	21.942	252.396	16551.973
31	199.768	18.922	5401.424	141.653	42.157	24.533	159.240	16186.434
32	197.509	38.093	5642.905	114.039	44.658	27.366	129.692	16020.408
33	161.873	-3.270	5276.762	130.297	36.409	23.605	101.210	13992.687
34	177.266	-16.986	5798.446	157.165	45.114	17.802	143.285	14941.617

Supplementary Table 3.5 Summary of Bayesian adjusted root ionomic phenotypic values for RDP1. Per line values represent the 50th percentile of the posterior probability distribution.

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
35	197.617	-8.955	5946.860	146.497	40.435	22.611	148.619	14971.974
36	162.099	14.445	5362.021	120.116	37.740	25.418	139.824	12749.163
37	155.929	24.984	5535.827	115.146	35.871	23.183	123.125	13269.379
38	190.163	10.118	5794.793	127.116	39.746	26.275	134.719	15704.518
39	197.993	-13.868	5750.356	164.567	44.478	21.775	129.744	15628.223
40	156.221	-14.896	5524.960	135.066	43.786	22.188	206.962	14083.653
41	189.783	-9.507	5981.383	135.498	37.326	26.610	169.452	16634.669
43	204.022	19.770	6067.567	144.612	42.143	28.384	236.982	16654.745
44	173.791	-10.092	5139.246	148.469	38.921	22.309	130.599	14607.361
45	207.861	-4.364	5348.361	140.728	49.727	24.884	249.950	13834.255
46	135.511	-20.180	5630.792	143.862	50.869	18.866	91.238	13504.507
48	170.568	4.276	5410.312	123.686	37.949	25.482	157.836	14715.719
49	179.052	7.627	5444.656	161.069	49.765	23.850	168.375	14875.103
50	151.105	-0.556	5430.705	143.060	44.408	21.839	128.756	13371.109
51	190.278	10.886	4999.492	113.229	40.511	23.558	143.905	13228.806
52	184.102	4.955	5689.565	119.619	42.408	26.478	146.191	14551.765
53	185.453	-0.291	5306.110	92.250	43.116	28.001	175.466	14351.932
54	184.069	18.769	5827.944	121.721	40.784	27.259	137.853	15661.060
55	156.035	-14.418	5356.012	112.807	39.629	20.898	141.885	12745.632
56	131.582	-0.250	5089.086	116.709	30.271	23.723	177.679	13895.536
57	194.473	9.218	5402.630	137.210	37.325	23.723	213.923	15766.717
58	195.268	11.496	5862.942	146.616	41.563	22.950	101.211	14307.703
59	187.849	-17.152	5673.648	120.102	40.982	24.238	137.021	14548.323
60	203.356	0.929	5539.719	112.330	39.917	27.208	134.748	13989.189
61	189.023	-6.646	5231.490	151.133	39.726	25.033	168.014	13615.704
62	211.723	16.254	5936.831	158.786	47.567	33.683	164.946	16752.385
63	205.150	23.239	5192.972	133.113	43.956	25.499	143.971	14230.260
64	189.908	13.172	5164.204	111.642	37.616	27.247	135.244	14092.726
65	226.783	1.262	6062.832	159.709	53.095	25.956	158.106	15981.381
66	211.909	8.059	5931.455	146.251	37.712	24.919	204.721	15494.874
67	185.496	23.915	5481.202	133.509	42.095	25.478	138.222	14677.133
68	254.156	11.303	6225.684	173.333	38.222	32.431	283.946	18922.467
69	143.657	-18.941	5969.775	139.288	44.898	21.284	81.222	13013.814
70	141.828	0.854	5571.469	134.190	42.356	24.463	147.703	13538.895
71	169.463	1.360	5638.176	109.023	30.417	21.739	130.638	15048.303

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
72	209.921	16.288	5826.088	134.765	34.534	28.929	219.862	18109.596
73	183.603	3.266	6006.054	153.343	50.510	25.231	145.733	16062.308
74	174.140	-11.738	5486.075	140.017	37.710	23.974	200.363	15376.574
75	149.308	-18.392	5440.081	129.984	51.726	19.718	103.055	14151.315
76	194.357	-1.427	5964.839	142.138	37.515	26.885	223.029	16046.180
77	186.311	-5.899	5388.632	118.251	38.357	22.344	133.489	13943.067
78	160.532	-9.838	5455.630	143.145	43.022	24.578	85.577	14298.487
79	168.435	22.003	5585.734	132.590	38.969	23.930	149.296	16120.190
80	170.628	50.425	5319.139	131.694	35.231	28.483	185.899	16305.010
81	129.109	-19.368	4976.774	137.151	35.574	18.325	102.816	11945.509
83	146.064	23.711	5084.053	98.829	28.998	22.115	97.647	12912.251
84	175.935	-1.731	5343.394	123.706	40.240	25.458	130.242	12321.576
85	185.024	8.867	4639.232	146.479	42.431	22.314	130.660	15414.709
86	128.333	-9.880	5335.933	127.895	40.440	21.873	145.601	14269.267
87	146.660	-4.390	5266.822	126.747	36.618	24.064	169.199	14754.851
88	151.152	-8.593	5278.759	141.987	41.098	23.771	141.595	13963.468
89	162.187	7.815	5478.104	134.996	45.584	23.805	156.986	13953.734
90	213.675	1.369	5679.915	124.641	43.129	24.248	177.125	15184.723
91	182.061	30.007	5152.316	115.429	44.459	24.638	160.256	15591.832
92	184.775	-3.037	5710.607	131.043	45.413	24.864	151.910	14367.823
93	198.581	-2.059	5069.653	87.483	42.756	26.495	157.842	13326.300
94	200.445	20.931	5087.750	127.214	45.021	25.771	171.330	15371.861
95	191.493	-21.398	4982.381	125.475	47.448	21.532	142.742	13703.916
96	213.024	0.728	5348.971	106.956	37.970	26.711	142.079	15237.356
97	208.698	10.819	5014.566	126.826	44.290	20.969	148.630	15532.339
98	193.803	22.991	6286.598	126.816	48.634	27.860	207.799	15998.859
99	231.964	-7.281	5708.221	126.613	49.507	27.163	181.896	16265.170
100	197.862	-17.746	5149.765	119.898	46.086	23.456	166.415	14595.155
101	194.209	27.435	5538.124	99.137	40.473	21.361	136.648	16522.697
102	236.838	23.097	5378.877	140.290	47.700	26.037	282.127	15955.625
103	200.479	36.386	4757.085	118.164	42.159	25.720	180.136	13960.374
104	224.279	-3.008	5385.812	117.048	46.695	28.827	156.579	15080.859
105	166.580	8.480	4723.084	115.603	42.408	21.712	99.366	12075.808
106	165.680	-2.755	4590.644	123.809	38.005	20.344	159.931	11078.718
107	201.980	-0.552	5403.110	143.100	54.344	23.194	145.245	14406.438

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
108	186.889	6.151	5201.287	110.669	42.090	23.068	164.579	13517.187
109	232.042	-3.744	5296.275	146.407	48.669	25.217	206.957	15856.493
110	240.163	-19.132	5334.055	132.186	44.575	23.780	252.859	14886.634
111	167.793	-3.152	5047.352	121.657	44.592	21.770	163.915	13636.301
112	230.873	3.868	5270.834	148.784	50.711	30.131	173.987	15521.875
113	255.372	3.797	5359.511	158.905	58.473	24.292	239.694	16630.782
114	230.582	11.144	6038.597	162.779	54.511	26.951	200.128	16243.115
115	208.376	10.324	4619.836	130.592	49.506	20.325	173.707	12621.671
116	225.717	11.060	5484.158	154.756	49.717	31.380	203.978	15873.628
117	242.450	12.106	5233.393	169.137	56.840	26.038	243.649	15285.157
118	189.121	15.141	5392.738	123.478	48.928	26.528	197.389	13807.284
119	199.709	3.587	4906.896	135.773	45.602	22.903	149.829	14347.211
120	211.797	-17.656	5250.336	139.338	48.911	24.260	151.581	14253.246
121	185.668	11.621	4801.316	129.798	47.675	22.767	168.553	14249.525
122	204.009	7.525	5328.398	123.994	45.733	23.244	262.003	14169.867
123	225.791	2.523	5498.188	129.192	44.515	23.510	216.042	15981.083
124	244.303	0.200	5465.965	156.256	54.331	28.096	240.178	14635.343
125	183.764	-3.577	5273.135	128.178	44.844	21.804	216.578	13691.835
126	189.943	8.138	4662.995	141.283	45.488	22.221	189.127	13478.478
127	169.186	6.301	5122.296	111.150	44.392	21.553	134.484	11822.864
128	213.252	17.629	5144.723	141.281	51.867	26.704	210.107	14969.660
129	185.990	1.798	5360.656	124.254	39.842	22.322	240.175	14950.494
130	228.740	3.925	5412.910	134.738	47.480	27.072	246.999	16353.467
131	191.380	7.719	5159.827	112.720	44.131	19.668	103.588	12983.554
132	200.916	2.656	4717.744	118.338	41.350	24.516	123.323	13369.082
133	195.966	43.258	5000.036	97.289	42.102	24.671	156.853	14096.284
134	183.032	-21.423	5092.474	134.354	57.723	23.365	169.120	13800.666
135	201.542	-3.450	5020.833	152.736	53.464	22.271	128.305	14096.011
136	236.732	-8.252	5377.040	127.490	41.220	23.164	173.146	15825.877
137	199.692	-8.773	5384.773	139.236	46.151	21.873	224.242	15359.837
138	221.633	-0.025	5235.868	145.824	48.647	21.379	137.743	15077.027
139	198.930	25.572	5181.755	123.673	45.681	24.578	163.652	13538.853
140	195.056	9.080	5333.756	123.440	51.587	27.755	192.530	13023.282
141	164.647	41.801	5608.750	118.260	40.406	20.856	146.977	12649.916
142	243.743	42.913	5739.082	174.573	54.493	26.351	186.808	18350.971

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
143	215.565	6.369	5171.680	121.850	48.643	25.477	192.048	15274.399
144	240.591	30.976	5298.913	155.182	56.365	27.458	240.120	16502.375
145	199.037	-2.779	4714.549	127.209	39.436	22.572	236.574	15108.725
146	288.128	5.083	5537.404	169.217	51.140	27.030	259.207	15686.940
147	217.203	19.223	5333.302	112.788	44.893	23.270	152.193	12940.347
148	186.396	-4.213	4904.492	130.472	43.832	22.908	178.194	15074.075
149	200.269	1.035	5262.535	114.385	42.475	27.167	140.471	14700.796
150	213.260	-12.228	5355.566	102.022	41.960	25.247	129.155	14278.386
151	262.555	26.733	5628.965	146.096	52.735	29.931	225.694	16412.427
152	224.731	-9.061	5375.324	130.915	45.990	25.054	142.018	15440.009
153	210.499	-3.704	4905.299	138.389	47.623	23.891	143.361	15897.950
154	206.224	-20.106	5468.701	121.416	48.447	23.431	196.308	14987.057
155	190.802	0.634	4971.699	129.360	49.568	21.004	138.787	12980.302
156	254.123	-2.474	5467.175	144.829	46.392	28.834	296.700	17703.486
157	222.845	1.840	4959.724	144.724	55.313	25.607	192.520	14848.677
158	209.486	20.508	5105.525	136.438	49.184	25.330	185.792	14616.503
159	227.657	8.672	5589.007	141.401	49.813	21.296	235.666	15654.966
160	274.792	3.629	5638.059	146.461	57.090	30.747	242.761	16737.641
161	258.774	21.438	5197.292	170.668	52.117	28.389	370.509	18497.973
162	254.705	9.887	5273.212	145.840	53.179	25.308	163.464	15322.820
163	245.092	0.852	5649.543	175.985	54.194	23.885	174.033	16479.206
164	214.037	-3.776	4807.953	113.842	43.665	22.580	230.904	12625.311
165	189.975	-3.917	4982.701	97.592	39.872	21.851	156.172	13883.542
166	252.229	11.903	5256.641	142.277	49.802	26.505	180.204	14885.920
167	206.323	19.359	5585.711	84.279	34.945	24.509	120.102	13967.537
168	228.807	20.026	5636.135	138.422	45.159	26.217	238.988	15336.637
169	194.541	9.825	4631.548	117.523	49.360	22.481	138.462	11634.757
170	219.020	8.929	5461.223	140.149	43.500	25.474	282.947	14032.183
171	198.962	0.375	5045.155	124.689	39.050	21.085	170.328	12487.914
172	170.552	32.702	4783.475	115.232	34.580	18.421	226.028	14859.253
173	170.321	66.389	3782.175	90.502	33.824	16.611	146.924	11401.056
174	130.454	50.735	3857.635	64.789	31.089	11.685	163.757	8912.662
175	119.720	39.796	3967.450	71.435	31.668	10.911	141.070	8564.209
176	134.858	57.231	3765.926	68.786	30.336	14.051	129.381	7311.828
177	147.456	91.783	3171.097	66.988	28.994	17.942	117.599	9078.070

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
178	109.045	49.020	2840.261	58.925	19.424	12.160	140.267	7711.707
179	123.322	52.136	3093.884	51.894	24.317	12.411	146.685	7129.680
180	102.150	54.537	2447.111	67.023	27.678	10.987	165.020	6288.553
181	137.462	63.236	2834.075	62.559	28.512	10.906	147.240	6095.881
182	108.804	62.622	4102.531	77.822	29.644	13.062	131.494	8173.822
183	127.677	68.851	3242.063	73.244	32.568	15.287	93.500	7923.092
184	90.940	35.164	3674.541	60.194	26.675	8.605	131.793	7231.760
185	116.749	73.247	3583.566	72.061	33.162	11.136	135.176	5785.760
186	130.725	61.540	3609.742	80.142	32.060	16.039	188.457	8515.547
187	98.565	84.364	3104.035	72.598	30.046	13.272	141.966	5904.275
188	157.728	89.769	3335.945	54.714	34.006	16.359	183.163	8172.069
189	137.651	62.656	2472.788	88.478	28.984	10.256	199.368	7444.623
190	93.752	54.583	3162.102	59.312	26.825	9.975	133.506	5968.882
191	117.817	45.092	3346.811	71.491	30.928	10.827	176.462	7276.054
192	146.657	68.231	3123.198	79.429	36.964	12.735	165.178	7098.523
193	90.108	63.159	2843.863	56.231	28.148	7.029	145.574	5535.547
194	106.744	51.064	2996.580	112.003	27.708	6.117	157.382	6555.723
195	135.873	72.617	4319.709	98.852	33.542	14.000	160.790	9444.682
196	112.957	47.827	3117.086	83.245	27.509	8.013	274.842	6143.185
197	105.773	51.897	3272.079	67.939	29.112	9.925	210.016	8807.540
198	108.072	58.476	3192.320	63.350	25.557	11.874	126.517	7309.448
199	139.187	65.608	3954.023	96.578	35.727	16.536	216.836	9760.804
200	99.437	42.589	3063.776	21.929	26.797	7.343	50.806	5660.713
201	93.080	56.721	3197.310	54.972	24.570	10.731	160.039	5872.832
202	111.544	68.060	3015.325	71.752	29.587	10.734	157.662	5246.054
203	117.787	51.408	3353.353	84.943	30.199	9.026	158.129	9095.157
204	108.272	51.690	3021.761	64.433	27.575	11.525	133.910	6471.399
205	127.111	43.923	3515.176	66.616	26.704	14.129	150.126	7300.043
206	132.079	60.786	4944.862	98.937	24.740	15.949	242.816	10586.970
207	122.978	57.488	3700.373	98.984	33.131	12.351	136.231	9376.768
208	177.615	58.120	3965.128	112.156	36.246	14.599	171.439	9022.835
209	140.860	57.776	4140.636	74.196	23.357	13.430	208.409	8324.163
211	145.059	74.534	3788.701	95.414	32.525	16.945	168.838	7376.416
212	176.943	70.532	3564.343	82.014	35.300	14.310	132.642	7595.548
213	103.333	58.699	3578.345	73.376	27.968	9.563	111.209	7374.630

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
214	96.278	56.871	3297.397	52.580	26.035	12.423	118.895	6277.357
215	98.949	44.249	3348.111	102.288	39.181	9.873	118.468	5154.210
216	133.978	98.311	3535.843	88.543	33.769	12.596	218.944	8169.065
217	99.889	58.488	2981.259	69.926	32.151	11.889	156.904	6603.109
218	132.367	72.091	3611.082	91.532	34.682	15.300	150.125	8915.474
219	112.421	46.479	2697.845	69.122	28.062	11.205	132.800	5261.734
220	117.390	39.431	3360.051	66.396	25.976	10.359	184.551	6920.877
221	142.096	35.001	3994.845	83.444	33.761	15.035	165.901	9051.188
222	127.814	29.584	3559.428	78.113	25.491	10.058	270.347	8764.499
223	113.475	32.227	3817.889	87.430	30.503	12.021	132.412	8911.974
224	115.698	39.757	3961.619	74.673	30.522	8.834	170.093	6536.636
225	95.857	48.407	3219.031	49.992	24.308	8.570	128.622	5810.930
226	86.875	54.128	3635.789	54.280	21.904	9.466	110.649	5552.508
227	114.868	49.938	3170.508	97.471	33.365	10.611	246.833	8445.547
228	116.468	39.273	3496.282	106.147	32.155	13.080	137.433	8447.636
229	116.953	52.949	3475.035	79.961	30.269	13.886	171.343	8131.082
230	117.215	55.114	2765.869	61.655	29.758	8.874	162.509	5925.757
231	158.524	56.692	3713.713	90.595	28.285	11.341	225.754	8162.400
232	141.874	91.402	3508.016	85.168	33.107	13.891	174.858	8089.602
233	127.529	71.084	2897.479	69.404	28.433	7.765	126.076	5325.375
234	113.793	45.057	3684.914	87.845	27.325	11.744	242.170	9120.663
235	149.532	43.329	3553.594	94.861	28.395	11.298	259.665	7922.510
236	106.467	37.419	2909.392	71.806	29.697	10.551	141.248	6615.672
237	121.329	53.776	3882.374	54.578	22.081	13.499	105.910	11485.614
238	94.426	57.399	4289.316	73.480	29.068	14.584	145.845	9295.095
239	130.444	72.949	2769.292	93.936	37.221	11.858	164.984	5490.616
240	139.863	60.603	4041.152	84.576	31.947	14.089	141.185	7199.534
241	97.038	47.740	3125.623	78.261	27.637	9.300	129.916	8413.477
242	163.887	61.286	3580.944	110.848	48.061	14.333	236.329	8368.501
243	96.284	46.360	2874.141	81.035	34.783	5.461	154.910	3985.359
244	141.177	65.165	2955.763	73.905	30.299	10.875	189.406	6608.096
245	130.409	52.635	3801.770	80.763	30.381	10.404	154.425	7173.735
246	81.829	50.543	3465.184	73.398	23.778	8.430	125.923	6648.400
247	102.874	58.978	2718.685	78.822	34.120	8.805	206.645	6450.646
248	101.845	70.170	2388.559	73.213	29.299	10.550	153.457	3292.285

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
249	103.711	54.141	3177.273	78.899	31.563	11.708	191.648	7894.465
250	121.960	64.422	3158.787	67.094	31.348	13.257	142.312	6993.823
251	168.439	56.434	3588.045	93.719	39.924	13.464	174.187	8741.180
252	120.178	59.885	2802.393	99.461	29.858	9.242	105.595	5251.163
253	119.664	73.952	2376.547	56.168	24.621	11.370	147.886	4077.167
254	76.150	47.790	2049.006	65.011	19.530	5.549	172.388	2769.427
255	112.978	52.291	2395.903	54.622	20.823	5.970	164.845	4648.755
256	113.214	47.209	2374.043	80.897	34.103	5.056	224.735	4699.639
257	156.704	68.975	2337.922	63.559	27.871	6.820	172.300	4319.878
258	149.664	68.989	2870.772	84.995	38.778	9.384	211.071	5321.298
259	96.088	69.386	1683.604	73.884	24.854	4.269	259.145	3664.482
260	138.429	65.412	1430.652	70.546	32.510	9.008	160.588	3475.717
261	93.585	67.778	2030.162	56.049	21.871	5.587	109.954	2398.887
262	109.474	61.057	359.506	56.101	27.243	1.616	134.246	834.579
263	107.509	80.932	1985.885	58.636	25.555	10.391	180.322	5952.410
264	110.856	64.830	1269.412	42.503	23.614	4.581	109.645	1119.599
265	106.511	67.811	1007.994	36.552	21.723	5.621	156.764	1696.281
266	103.728	63.659	2123.354	61.221	28.896	9.130	171.768	3385.851
267	125.467	68.951	1269.277	67.494	30.414	5.961	165.634	2599.868
268	117.373	72.845	913.140	43.190	26.413	3.201	137.683	-379.024
269	121.072	72.311	968.645	35.037	14.394	6.234	123.945	229.814
270	105.684	72.533	892.667	45.119	25.677	1.987	155.411	1169.573
271	94.316	83.031	1073.464	46.673	31.779	1.318	144.144	1260.061
272	103.259	85.861	1472.071	91.878	37.836	6.913	152.457	3334.590
273	105.775	62.564	1448.239	45.700	26.541	0.459	143.391	2055.571
274	75.348	60.760	722.718	38.933	23.830	-0.884	128.404	400.053
275	116.724	83.287	541.011	41.728	25.628	4.881	159.191	83.347
276	119.225	87.706	1136.167	114.651	23.404	6.861	274.175	3887.298
277	100.939	80.730	963.812	50.637	23.367	4.975	169.899	1403.570
278	117.274	87.179	1595.056	60.351	21.698	10.912	127.977	5139.950
279	71.048	72.125	575.244	51.840	24.502	-1.084	124.278	-542.070
280	85.802	68.330	1376.909	44.135	25.579	1.856	119.447	1265.680
281	129.214	84.158	1122.278	65.216	34.360	6.233	176.312	3598.186
282	131.057	90.187	985.091	71.710	34.779	6.681	237.616	3840.556
283	94.818	70.358	1348.558	50.908	30.664	5.407	147.839	1821.060

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
284	137.743	70.656	1524.247	83.217	28.490	7.180	289.564	4115.583
285	111.615	83.606	1439.522	34.621	21.366	10.346	184.069	3843.522
286	100.687	66.221	1395.408	53.731	23.810	4.004	136.033	2504.333
287	96.134	60.927	403.902	40.395	24.414	-2.071	142.523	-711.774
288	86.705	76.920	1042.430	50.262	24.061	3.499	182.098	1198.099
289	87.267	76.682	815.026	42.113	20.948	1.836	179.622	1027.640
290	139.365	100.611	2042.780	71.855	36.837	6.938	177.520	3953.078
291	102.915	87.701	1524.826	59.329	27.840	5.229	158.471	2272.760
292	119.068	57.930	966.348	58.564	27.955	2.475	164.913	2482.810
293	100.018	71.970	1079.649	66.384	29.034	1.978	202.302	1738.701
294	92.911	87.530	1361.070	66.538	22.110	2.759	156.402	920.748
295	99.935	74.583	590.261	37.251	20.338	0.691	190.268	-324.373
296	95.534	74.444	668.304	60.256	28.818	2.120	151.083	928.757
297	99.973	97.607	980.409	41.775	22.673	3.201	152.889	1706.041
298	164.179	82.657	1127.624	91.438	38.209	5.295	238.472	2246.669
299	121.180	82.358	2528.127	80.208	26.085	9.421	220.613	5161.870
300	99.790	70.233	1354.888	57.649	36.949	4.933	154.743	2047.371
301	117.560	57.173	1232.158	51.337	33.924	4.318	134.968	2400.676
302	107.621	96.722	757.015	36.428	27.110	4.950	110.461	1483.348
303	90.848	98.162	696.244	34.409	22.665	2.160	152.792	-158.320
304	124.589	77.331	1319.228	58.191	22.537	5.074	182.223	2896.554
305	99.501	61.612	1514.673	53.203	22.149	6.958	146.809	1779.174
306	98.886	60.580	964.960	47.762	21.797	-2.251	120.528	219.974
307	103.948	82.967	56.059	36.225	22.042	-0.228	175.764	-1713.313
308	104.175	72.460	1183.556	76.991	33.759	3.896	261.396	2197.734
309	90.212	68.159	1475.375	54.435	26.937	1.970	177.634	1897.589
310	93.676	102.952	1319.520	53.313	27.057	4.941	192.378	2008.656
311	126.256	99.727	961.368	59.370	31.682	7.385	159.831	2216.774
312	136.366	74.129	1007.015	79.634	30.755	3.682	190.176	2085.382
313	96.287	59.900	917.071	78.255	25.912	0.574	142.364	1099.159
314	129.359	75.457	1483.029	80.163	33.865	4.608	163.229	3601.420
315	63.175	47.211	1060.451	40.173	21.865	-2.270	91.570	2321.611
316	112.585	96.967	1083.595	66.672	29.613	-0.156	119.362	1507.581
317	99.813	64.119	1285.032	96.208	31.604	3.012	172.532	3419.609
318	112.511	56.325	1014.512	116.626	37.577	4.228	162.411	3749.992

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
319	103.931	76.128	-210.476	70.801	27.763	5.764	164.216	735.337
320	83.082	60.020	494.148	57.986	30.214	-0.275	114.276	208.263
321	128.872	64.307	1566.787	120.863	39.399	4.257	280.412	4694.919
322	103.406	73.040	898.266	57.545	26.158	-0.320	148.951	731.016
323	99.897	58.016	802.039	61.827	29.541	1.542	73.306	1495.139
324	119.267	70.659	1103.762	99.830	34.623	4.930	214.733	3682.085
325	123.084	84.512	1152.408	80.572	35.264	6.510	237.341	3889.314
326	113.272	75.001	1076.736	54.640	24.534	4.457	193.908	3431.849
327	107.478	56.689	1363.263	79.156	28.113	3.850	165.396	2685.437
328	133.261	76.383	982.913	95.034	38.453	3.631	222.590	1801.632
329	107.705	52.440	1066.850	59.336	23.537	1.252	114.174	2386.587
330	135.951	55.922	1513.462	80.470	28.890	4.273	116.842	2944.273
331	135.714	82.953	898.012	65.427	23.027	5.639	185.418	1938.133
332	115.650	99.972	1080.704	96.989	25.012	5.352	255.482	2524.792
333	120.985	93.237	1477.154	78.418	32.066	7.854	151.892	4134.754
334	118.892	67.418	1096.680	69.648	31.542	4.795	163.918	2633.176
335	122.952	70.294	1436.293	84.164	31.162	5.960	286.014	3050.348
336	81.253	83.689	788.255	42.165	23.197	1.086	100.552	1717.950
337	132.461	76.437	1512.225	76.258	28.000	2.872	184.319	2046.295
338	98.092	78.754	361.318	47.509	25.564	1.976	179.351	468.136
339	108.845	64.410	1232.282	65.025	25.529	0.025	180.577	1765.100
340	100.827	89.306	2062.235	53.559	28.876	3.605	151.447	3120.985
341	105.833	71.821	1956.648	62.238	24.034	4.355	130.248	4760.098
342	112.924	86.874	1632.266	69.866	33.312	4.932	202.763	2468.283
343	78.745	78.844	1038.478	47.446	24.743	2.183	192.916	904.397
344	179.399	91.665	2513.884	103.625	38.547	16.561	145.052	7939.302
345	103.417	65.436	408.922	77.138	27.651	3.071	150.707	1769.838
346	144.350	66.881	1998.316	97.100	37.514	8.579	102.019	4983.736
347	86.474	74.839	1504.914	53.912	25.141	1.864	114.052	1612.292
348	120.204	81.118	1719.876	75.469	22.370	6.553	209.986	4312.022
349	100.731	85.798	1853.992	89.120	25.483	3.916	153.508	2863.513
350	134.550	85.376	1954.407	85.843	39.279	10.106	235.798	4825.864
351	70.635	98.323	1575.953	43.640	17.907	7.109	116.266	3484.673
352	104.949	82.087	1980.392	73.551	32.738	6.459	161.788	4151.909
353	142.807	92.789	1633.113	69.402	22.056	6.194	136.494	4042.108

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
354	171.185	79.928	2288.460	105.384	36.089	11.335	239.172	8174.544
355	101.858	87.359	1647.346	62.420	26.373	5.499	205.095	3301.858
356	116.342	77.973	2116.708	87.944	28.453	5.824	258.602	3985.629
358	109.599	81.749	1550.940	74.168	27.246	7.943	162.490	3305.767
359	122.823	90.474	2010.648	77.492	29.362	4.732	184.926	4293.108
360	96.895	69.507	889.557	40.337	18.433	1.419	114.533	2752.277
361	101.655	86.412	2024.406	80.336	29.555	7.345	166.102	6038.276
362	107.331	72.673	2033.531	71.865	32.219	6.484	133.604	4077.237
363	86.625	89.626	1166.617	61.978	24.286	3.879	142.090	1852.220
364	132.948	79.892	1196.677	78.072	29.876	8.747	174.020	3650.684
365	115.062	82.962	2357.775	70.545	25.788	8.194	174.241	8275.518
366	109.011	106.000	1826.140	75.910	30.071	7.494	169.601	4302.054
367	95.483	92.800	1971.533	82.330	26.642	4.991	166.181	3470.721
368	118.679	97.928	1830.602	90.483	30.787	10.377	258.172	5949.967
369	73.642	46.525	1767.919	50.854	16.145	0.040	123.070	3553.019
370	121.608	70.597	2051.271	83.237	29.521	7.412	143.246	5086.275
371	108.738	67.894	1867.610	88.297	27.132	4.818	115.016	5118.982
372	95.580	78.679	1442.908	78.614	26.157	5.067	192.419	3770.567
373	104.522	73.391	1393.997	62.495	28.594	5.920	158.544	2997.675
374	114.339	87.689	1242.434	60.585	23.961	6.274	150.961	3895.342
375	111.695	76.037	2105.366	77.336	30.875	4.721	136.813	3899.104
376	97.231	78.762	2473.086	69.156	29.013	5.337	186.257	5842.415
377	145.671	94.870	2473.652	72.556	33.643	13.353	181.674	6258.354
378	110.404	67.616	2048.828	110.816	39.893	3.463	104.342	4550.771
379	127.532	80.328	2184.197	85.241	34.940	7.766	157.611	4419.428
380	138.370	93.116	1892.674	77.034	33.996	9.336	190.730	3929.750
381	132.034	86.050	2941.608	60.402	25.723	7.686	172.542	5729.031
383	63.700	82.590	1313.733	42.164	20.459	2.298	150.030	1619.574
384	135.853	79.099	2771.895	86.623	31.619	10.403	188.037	5520.970
385	96.571	72.912	1676.907	74.618	23.735	3.734	191.028	4619.308
386	102.041	80.212	2032.632	81.336	29.088	9.391	169.701	4952.719
387	77.604	78.853	2021.034	70.400	21.652	5.234	185.923	3988.306
388	91.086	77.249	1665.516	69.435	26.287	6.935	156.025	2248.647
389	146.901	108.927	2328.671	71.699	25.394	15.095	178.887	7157.419
390	120.041	99.249	2957.518	79.677	26.368	11.614	252.253	6565.657

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu	Fe
391	118.027	82.424	2203.428	73.049	27.159	8.184	168.773	4779.983
392	80.945	91.893	1440.764	57.183	25.357	3.242	119.830	2120.332
393	103.273	78.795	1888.493	74.181	21.568	5.000	187.206	4385.488
394	148.610	72.307	2340.339	111.112	33.623	12.753	180.876	6441.116
395	100.167	71.791	2513.334	81.855	32.677	7.806	203.571	4395.275
396	112.365	90.896	1808.843	72.769	31.655	7.804	163.776	2595.866
397	109.748	76.830	1503.790	58.286	22.337	7.687	181.053	3630.792
398	151.162	82.019	1836.385	92.659	31.753	9.363	305.080	5385.746
399	109.157	83.705	2164.242	76.082	39.150	8.046	159.741	4026.689
400	99.981	68.460	1834.983	78.258	27.613	4.468	219.718	3315.204

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
1	32.638	24084.038	25.539	1288.070	29.746	2.869	1505.012	-4.539	11657.027
2	32.505	33657.320	27.302	1354.964	36.485	0.544	1656.954	-4.291	9860.070
3	31.886	28941.361	23.288	1091.736	30.128	0.306	2051.481	-3.782	7869.454
4	28.885	28636.799	23.492	1093.099	29.661	0.121	1859.577	-3.121	8494.314
5	31.494	27798.636	19.405	1730.030	41.367	2.633	1276.821	-3.530	10412.086
6	31.723	32071.014	24.645	1032.882	35.565	0.211	1568.242	-3.612	9314.349
7	28.474	26478.562	26.566	1349.765	35.879	3.095	1154.755	-3.212	9332.230
8	32.838	21879.555	25.504	1428.474	34.784	-0.857	1340.759	-4.477	10241.172
9	32.059	19078.216	26.597	998.156	32.098	3.289	1367.405	-4.425	11652.644
10	32.182	23575.004	25.814	1229.717	34.043	1.543	1201.171	-4.805	10774.683
11	32.567	35880.827	24.302	1363.644	32.584	3.997	1648.843	-5.112	7781.255
12	30.887	21221.236	19.672	1479.699	42.873	1.404	1251.120	-3.783	7993.628
13	32.114	31429.225	23.350	1386.621	29.953	0.940	1463.285	-4.392	8322.161
14	32.260	22499.644	25.951	1590.482	39.113	4.912	1323.202	-4.732	10337.890
15	30.402	26460.739	26.600	1177.259	34.425	1.093	1310.606	-4.403	10029.708
16	37.130	30133.893	19.451	1633.098	44.150	-0.513	1529.556	-5.677	9786.162
17	29.632	23959.331	25.558	1772.422	46.080	1.437	1287.258	-4.069	9562.900
18	30.020	19152.179	23.668	1013.049	35.409	1.605	1632.867	-4.193	10532.047
19	28.807	23281.234	24.009	975.242	30.788	3.650	1235.038	-3.573	7658.183
20	28.526	27586.494	27.157	1179.024	33.558	0.042	1160.808	-3.733	11346.737
21	27.951	24779.599	23.589	1053.368	39.720	1.005	1240.403	-3.654	8140.774
22	29.160	28832.087	25.709	1338.144	28.479	-1.258	1298.048	-4.421	10154.972
23	30.456	20648.306	25.398	1515.579	38.420	-1.080	1272.089	-4.549	9173.851
24	31.452	25802.646	26.349	1412.704	28.619	-1.279	1130.309	-4.286	10966.233
25	30.829	24620.226	25.097	1463.780	34.223	-1.380	1287.666	-4.303	11964.900

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
26	31.025	26143.555	25.123	1527.494	32.726	-1.473	1209.675	-5.087	10888.014
27	32.595	25546.860	25.241	1547.421	34.717	-1.786	1597.686	-4.557	10669.535
28	30.014	34484.745	23.772	1313.784	30.078	2.102	1424.122	-4.085	8108.824
29	39.053	33431.874	23.978	1588.855	39.825	0.508	2059.456	-6.210	8947.843
30	27.230	20514.511	24.433	1348.212	29.472	3.690	1388.721	-3.174	9272.820
31	25.765	23144.345	27.580	1080.704	34.748	0.958	1277.983	-3.520	9488.647
32	24.774	30435.046	28.742	1197.299	26.269	1.771	1497.288	-3.235	11383.937
33	31.198	29641.614	23.980	1011.591	25.159	0.331	1587.850	-4.163	8465.996
34	28.300	27985.863	23.805	1494.697	31.241	-0.565	1548.609	-3.591	8290.265
35	28.604	24978.827	24.737	1396.725	26.172	0.002	1400.230	-3.914	8074.241
36	30.081	32469.889	26.413	1695.420	36.199	2.532	1446.125	-4.377	8652.061
37	30.456	29802.847	25.765	1820.013	35.106	-0.743	1311.025	-4.388	9111.801
38	30.240	27287.060	25.736	1329.647	31.198	2.842	1442.959	-3.989	9905.411
39	29.432	21215.901	24.842	1136.746	44.037	0.414	1967.785	-3.960	8700.827
40	30.504	23267.462	26.874	1501.113	27.658	1.700	1403.852	-3.961	10647.196
41	28.647	18381.794	25.401	1626.052	33.507	4.763	1383.542	-3.733	9960.517
43	31.671	21247.334	24.330	1490.106	50.254	3.387	1498.061	-4.257	9220.126
44	30.391	29358.498	24.180	1149.277	30.441	1.929	1666.741	-3.818	8191.379
45	30.251	22617.799	19.667	1603.434	43.206	-0.251	1573.522	-3.639	9326.259
46	30.615	22762.236	25.044	1325.529	40.356	-1.451	1255.417	-4.296	14052.412
48	28.146	23189.638	26.284	1481.865	26.568	0.744	1444.507	-3.521	10676.040
49	34.144	25122.617	24.512	1110.842	50.717	1.881	1757.650	-4.394	7747.392
50	30.402	31217.233	23.958	999.780	33.085	1.478	1426.943	-4.089	8080.723
51	26.519	25914.349	26.563	1096.010	28.892	2.338	1575.243	-3.710	9908.174
52	31.437	36203.123	27.058	1206.686	30.903	0.391	1439.108	-4.532	9628.605

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
53	28.877	21171.369	20.010	1475.925	38.464	-0.315	1562.126	-3.476	8567.686
54	27.119	28497.249	27.251	1176.199	29.906	-0.322	1530.838	-3.870	11383.943
55	27.183	26545.011	25.280	1125.426	32.817	2.044	1367.734	-3.846	10240.532
56	27.136	27617.472	27.113	953.546	31.319	1.072	1260.322	-3.535	8038.394
57	28.660	17190.644	24.000	1171.178	30.920	1.873	1338.267	-3.484	8554.238
58	31.778	35093.083	24.944	1731.191	38.171	2.314	1450.322	-4.362	9614.368
59	29.505	26340.874	25.938	1417.110	35.200	-0.775	1123.643	-3.620	11391.170
60	31.010	24919.536	27.098	1285.400	32.066	0.733	1471.330	-3.879	9589.395
61	32.100	33006.995	23.237	1354.519	37.362	0.514	1387.271	-4.251	8202.285
62	27.756	26947.880	28.192	1385.958	42.104	2.867	1243.950	-3.537	11308.808
63	27.176	23117.376	26.270	1268.881	30.897	1.605	1278.292	-3.575	10189.213
64	25.160	26790.048	28.250	1264.920	30.434	1.996	1487.578	-3.326	11066.170
65	29.589	18972.452	25.039	1283.412	25.394	-1.221	1282.895	-4.187	11919.896
66	31.765	30160.629	23.823	1550.715	39.260	-0.471	1593.064	-4.282	8917.415
67	26.813	23984.790	28.264	1423.508	34.874	1.045	1636.418	-3.421	9174.463
68	29.434	24159.306	26.755	1333.720	41.672	2.560	1295.034	-3.634	10684.446
69	32.835	26879.087	24.761	1136.154	39.926	-1.604	1398.028	-4.985	9621.833
70	26.800	26326.238	25.628	1425.109	32.740	1.512	1289.116	-3.420	9150.771
71	29.028	30746.111	23.450	1523.690	30.870	1.117	1613.629	-3.534	8938.488
72	28.780	23476.312	24.813	1652.515	42.671	3.818	1465.392	-3.782	8963.492
73	30.594	18925.480	25.012	1539.442	36.890	-1.491	1401.357	-4.355	10825.412
74	30.554	29529.460	24.308	1357.134	29.559	2.887	1550.817	-3.756	7279.936
75	28.096	24757.928	25.717	1410.643	26.943	-1.769	1407.240	-3.662	10665.557
76	33.669	27182.504	23.949	1573.714	39.038	1.823	1651.627	-4.881	8466.266
77	28.869	34483.202	23.382	1412.623	24.286	-0.549	1543.460	-3.299	8891.455

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
78	30.885	32348.910	23.547	971.269	37.814	0.157	1576.133	-3.666	7772.623
79	27.918	26453.008	27.573	1277.262	32.194	0.523	1406.133	-3.383	9659.379
80	21.885	21312.451	27.745	1613.082	31.080	1.268	1393.224	-2.255	9843.241
81	31.030	30396.755	24.078	1041.728	40.035	0.958	1567.144	-3.948	7350.214
83	27.073	28774.589	27.481	1066.598	27.309	0.101	1290.956	-3.191	8490.630
84	31.266	33082.443	26.192	1434.145	34.180	-1.117	1165.795	-3.993	10523.797
85	27.014	24940.353	25.065	1256.801	35.629	1.849	1522.470	-3.076	5825.970
86	28.934	28207.479	26.624	1336.992	40.378	1.486	1362.716	-3.839	11773.814
87	30.934	24018.060	25.439	1272.970	30.435	0.603	1311.543	-4.048	9543.605
88	28.357	36358.430	24.584	1326.182	33.313	0.489	1594.673	-3.501	9556.283
89	35.322	32912.380	25.851	1673.436	57.379	-1.262	1417.211	-5.343	11905.756
90	38.894	35410.518	22.957	1508.451	32.912	1.256	1752.003	-5.633	11135.705
91	27.659	29509.727	26.929	1402.943	29.800	2.515	1440.084	-4.174	12250.472
92	28.933	26270.088	25.618	1471.095	30.135	-1.872	1529.510	-3.985	10918.539
93	34.938	28557.949	20.140	1650.341	42.801	0.151	1662.053	-4.555	10562.985
94	37.437	31114.196	26.769	1545.057	30.616	2.090	1690.664	-5.486	11252.309
95	38.868	24557.249	24.713	1614.486	41.394	1.236	1330.546	-6.371	11104.263
96	37.991	23616.937	25.951	1900.046	40.647	-0.483	1590.207	-6.034	12103.452
97	34.828	28860.315	23.680	1258.581	33.422	-0.072	1372.328	-3.891	9766.931
98	33.673	33374.156	26.325	2037.056	35.065	0.371	1623.852	-5.039	12178.723
99	35.186	28087.461	25.145	1570.428	28.457	-0.950	1279.282	-4.498	12832.716
100	35.603	35404.499	24.949	1565.603	30.223	0.443	1541.306	-5.872	14521.569
101	28.021	29707.962	27.266	1608.012	30.093	0.405	1314.455	-4.159	14520.692
102	29.214	29377.178	24.182	2089.852	35.443	3.094	1346.712	-4.635	11267.382
103	23.337	26805.647	28.158	1453.355	37.489	1.809	1304.467	-3.520	11540.446

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
104	33.147	33984.688	26.900	1320.850	32.037	0.693	1251.100	-5.159	13629.669
105	33.955	38680.598	24.875	1394.379	32.329	-0.214	1665.037	-5.313	11858.056
106	32.687	39402.030	23.438	1422.748	35.428	0.263	1529.556	-5.773	9352.980
107	32.151	32796.871	26.923	1590.618	49.990	-1.242	1390.923	-5.708	13755.937
108	30.610	30789.633	26.331	1674.236	30.514	-0.575	1694.035	-5.376	12974.114
109	35.005	35694.117	22.953	1701.568	37.428	-0.970	1354.923	-5.598	12060.588
110	35.238	40009.725	23.164	1469.420	26.318	3.015	1499.671	-5.429	10682.091
111	33.650	31017.973	26.241	1723.946	55.166	-1.356	1417.132	-5.942	12603.774
112	33.829	27414.041	19.864	1587.168	36.539	-0.621	1455.981	-5.178	11625.469
113	28.914	31576.992	26.762	1669.319	41.110	2.042	1440.498	-4.481	13922.760
114	29.420	27218.452	28.166	1435.841	29.600	0.175	1591.952	-4.455	14064.992
115	28.056	27649.919	26.219	1755.812	29.858	2.172	1297.807	-4.355	12174.290
116	33.424	34551.818	26.314	1736.305	53.731	-1.419	1758.539	-5.774	12732.580
117	31.516	29419.724	24.946	1509.518	30.429	1.648	1493.276	-4.821	10553.934
118	29.156	29000.996	27.842	1690.061	29.711	3.792	1371.680	-4.424	12475.965
119	31.074	32025.195	24.738	1668.058	31.378	0.694	1139.488	-4.594	10699.600
120	36.910	25965.789	25.548	1557.193	53.455	-0.655	1369.291	-5.895	11678.278
121	32.617	25401.887	26.653	1301.942	32.949	1.752	1218.348	-4.955	11686.794
122	31.295	24887.611	26.756	1578.162	42.468	-0.226	1535.608	-5.018	12531.553
123	33.860	31192.989	25.455	1380.482	39.023	1.019	1670.022	-5.108	11984.499
124	35.542	30060.991	20.922	2058.862	46.668	-0.006	1656.545	-5.271	9587.813
125	32.952	34953.985	23.261	1248.070	39.627	-0.037	1609.694	-5.069	10553.253
126	27.013	32199.574	24.755	1401.234	35.466	0.901	1561.442	-3.972	9137.047
127	32.394	31308.774	24.862	1745.922	41.158	-0.667	1246.879	-5.679	12368.227
128	31.446	27919.986	27.011	1910.722	39.214	2.572	1349.942	-5.465	14392.753

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
129	35.092	25564.559	23.776	1296.200	34.064	1.772	1427.591	-6.193	10030.228
130	36.495	33845.898	24.887	1521.439	36.956	2.019	1492.679	-6.321	11293.008
131	30.602	29560.010	23.744	1346.626	32.763	0.332	1402.192	-4.841	10087.110
132	34.041	35147.790	23.351	1632.019	33.386	0.902	1601.799	-5.580	10356.907
133	27.457	35848.077	27.792	1540.959	26.064	0.875	1537.333	-4.247	13986.283
134	36.479	33802.493	24.999	1505.487	39.609	-0.437	1449.042	-7.291	13114.818
135	34.112	28078.432	25.912	1680.130	77.392	-1.588	1248.088	-5.948	11806.969
136	35.614	34743.021	23.933	1424.067	34.516	0.246	1559.800	-6.106	11775.014
137	35.427	28218.610	23.053	1510.230	38.096	1.389	1497.359	-5.109	11342.645
138	31.871	29729.437	23.706	1661.139	28.302	0.144	1474.376	-5.642	11310.857
139	36.110	31146.004	26.247	1861.286	33.886	-0.761	1486.289	-6.345	12168.634
140	31.241	32992.890	26.558	1820.913	31.984	0.462	1398.647	-4.598	11585.184
141	32.548	36328.794	25.645	1737.138	46.297	-0.419	1610.946	-4.747	9657.927
142	29.432	30823.055	24.537	1578.895	33.507	1.115	1356.647	-4.873	13692.789
143	28.544	27227.003	27.754	1488.514	37.159	3.248	1238.577	-4.859	11543.836
144	28.477	25436.634	27.919	1502.238	32.781	3.976	1047.051	-4.222	13926.919
145	33.741	37079.386	24.446	1533.386	41.345	0.616	1315.155	-5.343	11291.625
146	35.506	28681.673	24.055	1734.248	32.025	1.167	1587.970	-6.196	10928.911
147	28.587	29745.202	25.960	1677.986	30.610	0.222	1359.544	-4.744	13143.200
148	36.158	34750.194	23.641	1446.032	28.062	1.494	1524.939	-6.589	10436.012
149	35.366	37447.403	24.842	1554.847	30.388	0.365	1529.549	-6.361	13852.928
150	31.568	26745.344	24.778	1648.087	31.560	-1.176	1425.552	-5.444	12798.649
151	34.256	28933.378	27.127	1938.143	42.209	4.711	1656.606	-6.108	13897.106
152	37.941	32038.381	24.105	1379.658	30.678	1.404	2031.887	-6.129	10756.374
153	32.796	33434.214	24.212	1410.069	31.085	2.040	1578.706	-5.461	10454.911

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
154	38.183	37402.825	25.356	1870.469	49.879	1.034	2021.159	-7.292	12974.500
155	33.071	31035.812	25.502	1728.575	38.750	0.514	1423.484	-5.883	11997.918
156	37.148	34759.963	24.023	1569.407	41.726	3.347	1302.474	-5.974	11871.442
157	31.689	30829.883	27.322	1619.482	44.492	0.881	1413.793	-5.269	12187.927
158	32.512	36729.788	27.404	1649.817	36.746	2.009	1389.819	-5.701	13318.722
159	31.966	26610.081	24.035	1411.427	34.528	0.992	1533.029	-5.521	10955.429
160	37.337	31520.314	20.210	2035.717	45.747	-1.001	1462.540	-5.917	12919.088
161	30.310	29242.959	25.158	1289.544	37.996	3.722	1363.588	-4.267	11941.498
162	35.815	27956.425	23.434	1649.343	31.617	1.556	1606.313	-5.726	11061.284
163	36.119	28561.791	23.368	1596.699	36.963	0.764	1589.349	-6.315	12470.783
164	32.550	34302.179	25.748	1449.548	28.619	0.156	1613.872	-5.808	11661.398
165	32.242	27711.126	25.284	1647.037	27.367	0.013	1628.445	-5.652	12336.708
166	30.547	21714.005	25.915	1320.610	34.354	2.453	1345.537	-4.562	10395.004
167	28.858	31498.062	26.272	1703.402	27.125	-0.147	1484.841	-4.944	13942.843
168	31.650	32492.309	27.681	1672.512	38.905	4.752	1643.336	-5.169	11026.055
169	30.843	34677.820	26.825	1879.728	34.352	0.772	1490.989	-5.553	12960.037
170	35.101	30220.182	25.218	1604.198	42.882	1.146	1551.720	-6.249	11127.024
171	33.566	35460.033	24.449	1656.787	31.627	1.291	1546.654	-5.451	10746.929
172	21.929	29348.559	25.877	1129.550	30.747	2.457	1235.253	-2.006	10217.754
173	14.622	23574.926	28.741	1224.900	29.189	3.356	1341.059	-0.935	11743.679
174	7.156	20617.626	27.620	1270.006	15.100	0.745	807.947	0.898	9757.388
175	6.064	22648.021	27.126	1372.470	22.310	1.155	846.906	0.940	9089.203
176	9.295	25637.331	27.446	1657.832	20.954	1.420	968.394	0.261	10001.389
177	5.904	18280.248	29.151	1013.039	17.192	2.540	1194.681	1.181	9400.172
178	2.865	19580.866	26.353	798.113	13.869	2.750	1123.150	2.041	7180.471

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
179	2.682	21793.691	28.634	1236.286	22.657	4.099	807.162	1.814	10439.948
180	4.811	24388.012	28.381	1265.452	15.579	4.391	565.638	1.678	9480.511
181	1.460	22713.055	28.649	1196.104	12.842	5.889	1144.821	1.582	9798.743
182	3.653	22721.458	28.268	953.242	20.351	1.971	687.977	1.560	8533.491
183	-0.487	12979.617	28.490	1035.293	38.742	0.445	798.552	2.759	10403.708
184	-1.684	16424.808	28.452	957.625	18.920	2.328	919.977	2.651	9193.977
185	-1.499	16151.056	28.426	1083.460	19.766	1.035	795.629	2.331	8404.525
186	3.571	19204.994	28.889	979.992	41.446	3.343	720.033	2.031	10396.322
187	-1.173	14867.984	28.655	1068.820	14.267	1.923	904.819	2.440	8450.616
188	-2.746	22633.667	28.806	919.565	16.094	1.591	743.552	2.890	9727.299
189	1.280	29756.944	26.320	1177.718	22.720	2.762	1308.675	2.247	9267.421
190	-2.962	17001.141	28.685	1188.576	21.578	1.378	676.950	3.345	9275.494
191	-1.398	17229.156	23.134	1138.795	23.463	1.921	1067.057	3.152	9135.669
192	-1.118	20531.783	29.061	1025.310	20.040	2.862	1035.596	2.464	9527.601
193	-1.355	21068.380	27.390	1166.450	16.470	0.510	919.597	2.690	8396.828
194	-3.380	18501.778	28.396	1567.361	22.569	3.262	1137.469	3.371	6523.138
195	-1.825	15875.781	28.627	1437.353	25.809	0.971	965.046	2.516	10917.385
196	-1.276	22234.211	25.856	1094.856	11.764	3.439	823.685	3.083	6405.723
197	-0.992	11678.516	28.791	1499.106	19.236	3.010	840.091	2.669	11891.026
198	1.464	17143.974	28.285	1360.249	31.383	0.901	1110.777	2.458	8854.417
199	-0.063	19733.363	28.361	1120.977	14.582	1.086	834.982	2.652	9958.703
200	-1.802	25406.474	25.872	980.496	18.733	2.914	1327.293	3.152	6465.443
201	-1.885	20676.722	28.465	1235.215	15.549	2.322	812.732	2.778	9440.072
202	-3.208	16444.543	28.359	1140.228	39.998	1.091	816.145	2.724	9635.146
203	-1.875	19813.114	26.192	1042.624	17.960	4.687	1017.315	3.036	7170.037

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
204	1.661	20445.794	28.451	1257.921	13.851	2.709	888.525	2.259	9539.576
205	0.477	17836.308	27.929	999.939	23.951	2.555	918.913	2.256	9645.873
206	1.221	16641.819	29.379	1160.896	18.561	5.444	1024.471	2.889	8695.098
207	-0.041	19648.077	26.554	805.508	15.663	2.790	885.186	2.682	7870.272
208	-2.008	19375.106	26.834	1028.674	18.253	4.022	932.294	3.109	7426.195
209	0.634	23569.735	26.836	1036.935	22.197	2.642	1194.672	2.180	8303.126
211	-2.424	15478.012	29.063	1187.024	31.489	2.903	787.236	2.871	10835.188
212	-6.311	16960.461	28.375	1139.199	14.461	1.484	764.624	3.197	10485.322
213	-4.172	19496.027	28.180	1013.297	14.453	1.590	628.933	3.107	10958.578
214	-2.103	20027.728	28.110	1354.271	22.007	0.945	1202.154	2.731	9237.074
215	-0.144	19956.610	27.712	1687.220	46.917	0.503	1006.702	2.321	10485.280
216	-3.003	12815.273	30.480	1091.231	26.428	4.318	1065.710	3.197	8079.038
217	-2.774	14325.963	28.475	1153.318	13.355	1.981	726.264	3.488	7463.669
218	-2.501	18215.326	28.488	1223.843	24.563	3.643	553.366	2.872	11311.317
219	3.620	17149.196	28.271	1025.292	23.339	2.256	748.869	2.228	7779.533
220	3.178	21254.150	28.623	1016.679	20.083	1.758	1105.406	2.307	9972.074
221	7.199	9679.557	22.101	1584.910	22.817	1.731	937.967	2.052	9074.056
222	7.001	24896.072	25.928	1136.190	25.189	3.524	1119.890	1.403	7967.104
223	9.033	21789.740	27.198	1060.130	28.128	0.532	866.399	1.200	10111.989
224	1.063	16534.819	28.155	1056.822	22.390	4.111	723.355	2.577	9614.592
225	1.503	21472.303	28.700	956.739	18.564	2.967	893.144	2.736	8892.424
226	1.145	12964.744	27.801	1200.312	23.585	1.539	757.399	2.613	7500.072
227	1.087	17771.597	28.358	1210.367	21.603	4.325	850.810	2.881	8480.724
228	7.524	21909.341	25.968	817.463	29.890	3.841	1220.086	1.604	7422.231
229	6.158	23719.043	27.629	1077.651	30.896	3.365	779.543	1.719	9946.615
230	-2.214	23737.136	28.614	1040.416	11.581	4.247	730.461	3.040	9198.864
231	3.656	24187.973	25.959	1088.545	21.286	2.525	752.362	2.593	8816.925
232	-1.968	16466.836	29.349	1141.215	24.160	3.449	791.967	2.773	10238.203
233	0.859	21199.371	29.009	1224.136	18.957	2.449	1187.495	2.663	8328.701

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
234	7.551	21663.480	26.341	1154.873	29.171	1.949	990.564	1.483	7997.530
235	2.934	22967.688	25.703	756.351	18.169	3.936	1167.343	2.317	6983.216
236	2.024	23717.325	28.483	1019.742	27.156	1.933	821.576	2.452	8680.443
237	1.043	14639.939	28.554	1050.464	21.903	2.410	928.210	2.578	10482.980
238	0.568	13036.860	28.604	847.699	16.123	1.371	1046.325	2.775	9353.445
239	-5.694	17358.467	28.246	1398.887	14.303	1.363	460.841	3.503	8514.052
240	-3.199	17579.491	28.019	1615.002	26.064	1.827	758.639	2.756	9562.303
241	-3.793	18333.321	26.087	1013.741	14.928	2.223	716.181	3.042	8009.347
242	-5.763	13627.465	28.994	1187.475	19.998	0.981	807.226	3.114	10052.267
243	-2.064	16630.103	28.936	1046.694	23.900	2.644	769.621	2.834	7164.485
244	-4.373	17356.040	28.679	1351.048	16.136	3.560	578.325	3.472	7625.181
245	-3.357	13512.423	28.711	1060.713	24.854	3.840	959.440	2.951	7855.828
246	-3.832	23127.603	26.580	853.736	18.471	3.680	775.917	3.455	8138.422
247	-6.012	15836.789	29.065	1067.676	15.935	3.889	781.803	3.518	8051.686
248	-9.751	20145.554	29.792	1265.535	11.356	4.180	1049.947	4.272	9015.303
249	-2.743	19644.950	28.374	1367.239	16.391	4.362	852.824	3.182	8804.538
250	-6.110	15776.450	28.864	980.042	15.628	2.698	640.245	3.720	9576.745
251	-3.923	14222.396	28.226	1221.261	15.371	2.661	706.485	3.177	7634.901
252	-4.952	14933.984	26.644	1330.603	12.499	2.958	924.975	3.661	6738.239
253	-15.067	16080.521	29.692	1148.188	14.036	3.352	644.643	6.015	7259.353
254	-2.300	18640.084	26.237	979.151	18.518	1.867	653.387	3.474	5756.786
255	-6.273	23115.350	26.195	811.449	14.189	3.492	1088.209	4.371	6117.537
256	-7.881	17287.296	29.409	927.992	19.509	5.131	622.928	4.572	8873.001
257	-10.460	20101.857	29.186	1081.162	14.241	4.956	658.168	4.973	9479.976
258	-10.505	11116.847	28.928	764.793	11.785	2.826	691.649	5.004	9042.928
259	-18.869	11124.568	29.515	911.586	14.956	4.768	455.747	7.028	5421.097
260	-18.557	11504.251	24.285	1579.428	28.122	3.801	675.985	6.714	6577.957
261	-13.842	15889.732	27.155	927.228	9.880	5.580	967.515	5.288	6139.051
262	-23.050	15543.449	27.147	1105.936	8.136	5.029	606.686	7.705	4699.748
263	-14.401	13129.063	30.055	1165.287	12.204	7.656	908.143	5.278	12057.518
264	-16.560	16990.289	29.221	1201.398	15.737	4.073	809.343	6.304	7736.584

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
265	-18.512	12131.432	29.777	1112.914	18.933	5.137	647.421	6.469	8215.031
266	-16.374	10868.126	29.562	992.852	17.393	3.408	991.941	5.554	7106.353
267	-20.432	13897.835	30.198	804.483	10.535	4.287	720.355	6.400	6860.076
268	-26.188	10519.310	30.318	944.521	16.835	4.917	661.342	7.492	5678.177
269	-19.167	15065.325	27.615	895.240	13.847	4.874	828.554	6.678	4875.001
270	-27.768	8098.941	30.005	825.255	16.052	6.169	775.833	8.574	7621.172
271	-24.133	9295.722	30.087	1052.849	9.732	6.638	878.801	7.536	6996.224
272	-24.357	6767.813	30.437	1009.241	36.901	3.316	621.399	7.976	6314.386
273	-26.790	11064.216	30.118	1036.791	6.652	3.630	825.920	8.711	6913.823
274	-24.679	5942.971	29.429	915.593	10.811	3.032	443.951	8.642	6170.996
275	-25.152	12536.628	30.488	1027.696	11.491	5.665	887.367	7.146	6612.990
276	-23.668	5481.643	28.365	588.667	7.412	8.233	704.193	8.291	6442.714
277	-26.574	10229.887	30.605	798.346	10.001	6.195	450.911	8.318	7549.606
278	-21.104	14178.491	30.088	881.967	22.522	5.517	583.128	7.286	6986.388
279	-29.372	5452.948	29.747	781.393	9.745	4.207	426.855	8.791	5620.328
280	-24.759	7880.316	30.071	778.790	19.358	2.956	696.883	8.083	7404.117
281	-20.790	8727.149	30.523	822.270	13.927	5.165	453.362	7.110	7573.003
282	-24.746	8378.099	30.910	1202.692	13.688	6.933	662.707	7.609	8881.407
283	-23.452	8406.589	30.182	922.201	18.507	3.715	710.227	7.160	5756.786
284	-19.368	8108.564	27.623	1360.931	10.268	5.015	601.163	6.828	5140.763
285	-19.696	11042.726	30.354	891.113	11.139	3.261	704.848	6.779	8652.573
286	-20.784	14376.246	28.977	1075.987	7.909	2.609	697.528	6.878	7162.816
287	-27.737	14450.813	29.255	927.347	11.754	4.054	409.659	8.073	6162.570
288	-25.353	8447.244	30.689	1162.208	14.789	7.085	566.155	7.861	6653.860
289	-26.860	5392.080	30.223	1037.274	13.860	5.398	587.173	8.097	7242.991

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
290	-21.022	7557.675	30.794	1060.315	17.530	5.130	923.232	6.341	7810.796
291	-21.092	17112.762	30.542	994.411	19.694	4.331	820.101	6.738	8469.353
292	-22.130	12373.877	29.752	707.122	8.468	5.332	769.599	6.998	6693.087
293	-23.473	14202.035	29.791	963.051	9.987	4.656	796.050	7.640	5557.607
294	-24.153	4575.596	30.292	948.556	14.469	3.216	624.904	7.968	5858.674
295	-24.944	7777.354	30.283	1039.778	5.630	8.061	510.909	7.682	8309.591
296	-24.310	11964.678	30.289	1063.537	8.144	6.684	815.436	7.452	8507.203
297	-26.627	6405.148	30.963	638.384	7.422	6.026	654.799	8.031	7968.497
298	-21.174	14916.816	27.837	1249.380	15.909	6.811	593.548	6.689	6684.721
299	-18.949	5785.516	27.959	1071.159	9.771	3.900	1072.792	6.482	6390.587
300	-19.206	13021.727	30.288	1145.147	12.832	3.000	846.829	6.695	8737.660
301	-21.654	4623.243	29.800	816.510	15.313	5.068	638.935	7.203	7277.003
302	-26.725	5176.346	30.707	801.036	6.296	3.905	763.371	8.028	6203.339
303	-29.369	10619.077	30.629	1021.692	7.081	6.488	625.195	8.183	6918.819
304	-20.467	13049.732	27.616	884.925	11.857	3.072	743.440	7.145	6124.911
305	-18.584	13991.598	30.141	840.633	18.831	3.342	1100.910	6.601	9024.688
306	-24.254	10163.751	30.044	595.473	8.038	4.996	590.636	7.608	8460.492
307	-25.580	15692.267	29.993	970.025	12.521	6.308	519.760	7.472	7940.629
308	-21.306	12032.036	29.226	1059.566	13.663	2.717	641.126	6.690	8654.823
309	-24.687	6065.870	29.090	1128.579	8.260	2.443	385.742	7.593	6574.533
310	-26.648	9727.774	30.075	1241.450	6.752	2.127	613.748	7.679	8291.213
311	-24.355	7004.026	30.844	786.295	10.505	5.128	658.952	7.504	6377.448
312	-20.171	8860.767	28.233	724.774	21.077	4.449	1175.275	7.029	5365.506
313	-19.932	16255.466	27.348	778.949	17.519	3.840	619.551	7.213	5841.578
314	-19.914	12154.640	28.170	647.873	15.197	5.590	1197.147	6.840	6425.629

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
315	-20.308	15729.361	27.038	831.117	8.269	2.769	792.760	6.708	5057.133
316	-26.129	11366.156	28.185	749.874	10.765	5.417	786.651	8.078	5323.722
317	-20.720	11236.075	27.960	712.510	8.022	6.315	988.371	7.040	6652.557
318	-20.310	9287.348	27.426	814.279	10.188	6.166	873.316	7.150	5566.742
319	-20.895	10077.449	27.559	1423.104	22.590	5.857	552.929	6.915	4951.003
320	-23.308	14618.383	27.264	876.985	24.090	6.616	578.503	6.990	4189.610
321	-17.992	10875.832	27.892	757.728	24.545	6.073	916.328	6.449	7036.139
322	-27.778	8395.780	28.155	600.487	7.490	5.656	1225.407	8.076	4744.263
323	-23.632	17868.181	28.347	806.099	18.033	4.967	1180.880	7.357	5578.538
324	-19.956	10216.751	27.705	771.892	19.018	5.653	765.290	7.634	7422.495
325	-16.532	19020.175	28.012	1063.020	29.220	5.265	529.026	6.388	6736.157
326	-23.606	7193.114	28.097	732.355	7.861	12.532	1077.294	7.782	5335.549
327	-19.912	17939.913	28.011	1066.603	12.033	5.855	777.694	6.767	6433.020
328	-21.207	12113.415	28.062	897.156	17.580	5.639	806.908	7.280	5399.111
329	-19.063	12308.541	27.197	647.799	6.852	7.115	796.784	6.970	5154.492
330	-19.858	11613.592	27.039	1028.078	13.499	3.998	737.381	6.383	6438.488
331	-27.013	9371.861	28.439	573.426	11.701	9.112	774.934	8.309	5078.328
332	-28.546	2492.532	30.190	958.723	18.728	7.388	509.648	8.426	5228.275
333	-18.858	5233.661	30.295	716.680	10.381	4.402	504.416	6.931	7635.271
334	-21.700	8929.287	29.805	914.757	9.365	7.152	477.866	7.097	7960.897
335	-19.516	9261.437	29.851	1267.457	30.090	3.132	641.596	6.752	8306.971
336	-21.847	13978.720	28.108	482.160	7.454	9.452	878.856	7.653	5586.290
337	-16.969	11590.787	27.378	681.013	22.716	3.804	925.590	6.366	6129.320
338	-16.955	9858.325	30.149	786.390	10.055	6.418	583.736	6.732	5441.061
339	-15.452	17587.600	27.078	865.622	21.313	3.170	781.382	6.392	5176.763

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
340	-18.938	26335.343	29.897	978.541	10.079	3.644	694.351	6.509	6151.214
341	-12.748	17193.396	27.710	908.711	18.319	4.255	1146.072	5.747	6451.533
342	-17.434	8047.277	29.631	882.464	19.731	2.336	643.418	6.654	8157.378
343	-19.474	15700.441	29.580	1089.859	13.319	3.476	438.357	6.882	7797.288
344	-15.651	4213.356	29.817	983.738	25.324	3.728	949.209	6.106	8687.597
345	-15.514	17378.059	27.280	597.022	11.823	5.479	983.563	6.128	5649.611
346	-13.550	16301.827	27.083	810.345	27.296	4.489	676.901	5.372	7410.817
347	-16.823	14222.081	29.213	1096.930	13.517	2.062	668.391	6.095	7098.484
348	-11.801	14591.070	27.493	1166.627	20.350	3.654	960.398	5.831	7163.421
349	-18.040	11782.832	27.371	618.798	15.356	4.054	698.376	6.409	6077.618
350	-14.470	14848.297	29.333	1164.983	25.529	3.243	576.542	5.394	11418.033
351	-18.211	5672.602	30.402	640.178	10.328	4.270	882.203	5.965	6264.466
352	-16.721	13631.435	29.332	1106.880	11.156	2.665	748.030	5.941	8280.809
353	-16.385	14555.878	28.027	715.888	23.252	5.675	997.612	5.726	6744.331
354	-10.471	13206.481	27.293	1556.504	23.476	4.141	1025.815	4.879	9813.888
355	-15.801	16103.055	30.558	1432.695	14.261	5.747	1284.842	5.691	9287.254
356	-10.625	17885.265	27.048	1136.654	28.597	4.118	1027.380	4.999	7446.072
358	-17.660	18044.293	29.262	1077.482	9.999	3.375	721.619	6.008	8969.841
359	-17.927	15675.628	27.959	797.118	19.561	7.346	1299.171	6.199	7326.557
360	-14.105	21037.710	26.858	1112.648	16.803	4.124	1139.181	5.709	5260.994
361	-15.621	10954.200	30.294	912.699	12.224	4.941	769.569	6.131	8690.508
362	-10.112	18258.878	28.832	1110.908	13.873	2.696	720.629	5.108	8047.608
363	-14.983	14447.004	29.916	1015.940	14.305	5.791	743.671	5.780	7269.739
364	-16.493	14266.062	30.109	1268.233	15.482	3.853	849.333	6.042	8361.280
365	-9.545	16521.288	30.433	1064.685	17.105	4.514	853.568	4.772	11198.487

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
366	-16.279	16699.852	30.582	1234.757	22.575	6.038	944.580	6.289	9032.511
367	-15.919	17947.716	30.074	1267.201	15.408	4.663	691.271	6.323	10074.685
368	-13.618	15787.924	29.890	1188.716	20.181	5.256	429.110	5.918	9309.551
369	-8.558	22727.676	26.082	929.892	10.062	4.482	1091.733	4.561	6839.595
370	-7.843	13914.029	27.712	752.510	20.453	6.207	800.989	5.446	7443.302
371	-12.149	15133.148	26.934	758.748	15.039	4.330	896.293	5.579	6825.338
372	-16.642	14931.088	27.483	969.143	15.809	6.620	858.917	5.879	6520.652
373	-15.701	17282.272	23.593	1264.653	20.735	5.645	832.923	5.678	7567.531
374	-17.103	13820.305	29.655	1063.159	18.470	6.244	643.248	5.807	9186.418
375	-16.694	14282.009	29.225	1016.987	11.684	3.116	650.537	6.487	9697.036
376	-15.760	6744.850	29.687	955.867	22.328	4.023	894.124	6.963	8751.616
377	-15.280	14682.871	29.563	1250.818	16.591	2.024	605.079	5.669	10695.553
378	-13.037	12452.501	27.876	1145.818	35.673	4.428	1097.997	5.979	7371.601
379	-15.572	12366.491	28.400	1268.453	15.204	2.158	559.304	5.805	10148.054
380	-13.778	12306.027	30.069	1105.550	17.386	7.742	966.785	5.289	8603.743
381	-14.680	16669.847	29.716	1253.413	31.415	3.510	937.403	5.507	10525.307
383	-17.909	14719.941	29.307	1182.756	17.831	4.278	1018.190	5.936	8791.818
384	-16.036	10845.354	29.943	1244.437	16.982	2.438	1079.461	5.719	9975.060
385	-11.029	13395.544	26.710	679.927	15.733	4.910	790.444	5.303	5928.541
386	-15.390	13219.975	29.934	945.863	12.337	3.039	918.164	5.740	8976.698
387	-13.117	17849.703	29.757	1764.378	18.640	3.781	909.628	5.318	9602.972
388	-13.166	13015.000	29.721	970.623	19.218	3.845	754.618	5.651	6909.203
389	-16.842	11941.121	30.677	1225.305	15.771	2.881	956.809	6.098	10728.519
390	-14.482	13759.793	30.222	1189.059	28.925	3.453	849.037	5.673	8968.801
391	-12.876	11461.126	29.019	1243.582	20.560	2.677	825.347	4.894	9419.250

Supplementary Table 3.5 (continued).

NSFTV_ID	I	K	Li	Mg	Mn	Mo	Na	Ni	P
392	-22.962	15974.617	29.599	959.455	14.422	3.036	669.386	6.563	8122.535
393	-13.375	19644.575	27.238	1151.187	25.573	5.079	763.285	5.528	7253.193
394	-14.484	12933.446	28.943	1431.563	13.758	2.558	810.880	4.686	9956.441
395	-11.378	15165.487	28.816	1146.822	18.248	1.949	951.734	4.584	9181.219
396	-22.917	8965.856	30.295	840.138	9.142	3.303	917.316	7.111	7017.069
397	-13.190	25791.943	29.043	1197.726	21.649	3.255	513.572	5.239	8999.584
398	-11.568	16096.522	27.668	1049.095	18.856	3.046	921.011	4.964	7592.715
399	-18.940	13036.479	29.787	913.396	12.082	2.348	740.757	5.733	8065.630
400	-12.832	19006.595	26.888	1002.472	19.895	2.597	918.877	5.260	6842.589

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
1	21.878	-47.638	4403.218	258.106	524.148	9.019	929.785	0.140
2	24.567	-42.547	4344.442	206.527	521.698	8.464	1100.864	0.106
3	19.689	-45.623	5574.667	231.629	418.926	7.989	804.760	0.121
4	20.000	-44.040	4329.200	190.203	400.650	8.413	935.345	0.079
5	22.476	-49.075	4492.683	243.817	467.838	8.835	1030.071	0.125
6	27.965	-43.098	4846.063	232.182	462.195	8.999	859.249	0.112
7	21.868	-33.820	4135.848	255.962	606.436	8.606	945.807	0.026
8	21.994	-49.239	3836.144	233.061	483.155	9.758	961.998	0.116
9	23.061	-46.581	4028.336	283.311	515.131	9.829	995.646	0.128
10	23.845	-49.403	4233.431	264.556	550.302	8.865	1112.353	0.042
11	25.787	-42.276	5960.176	316.232	507.653	8.027	884.227	0.072
12	17.200	-48.761	4767.644	255.517	435.476	8.830	919.484	0.117
13	28.410	-45.888	5554.821	241.488	575.181	8.442	903.005	0.177
14	30.299	-43.270	5463.038	316.119	453.254	11.178	914.814	0.092
15	25.514	-44.414	4354.328	246.388	502.648	8.685	1069.821	0.041
16	26.479	-69.064	6073.891	277.838	493.832	9.792	1029.683	0.105
17	24.863	-36.895	6364.205	296.846	514.311	10.242	962.491	0.051
18	33.700	-44.771	4756.995	343.111	526.907	11.281	793.438	0.116
19	21.744	-37.841	4273.098	210.184	411.370	8.423	765.169	0.150
20	22.021	-45.921	3882.289	230.803	528.830	9.056	1079.559	0.074
21	27.422	-35.563	4804.931	329.656	509.448	8.765	934.993	0.057
22	18.980	-44.055	4330.871	211.494	391.317	8.002	969.237	0.159
23	23.148	-49.647	4980.802	279.887	510.583	9.535	1134.846	0.147
24	21.015	-46.131	4244.691	271.438	536.701	8.802	948.473	0.104
25	28.368	-47.065	4365.375	263.513	526.025	10.676	943.584	0.088
26	22.465	-50.897	4725.750	239.920	517.674	8.702	933.449	0.099
27	25.616	-50.881	4827.800	273.927	463.985	11.344	1065.297	0.077
28	25.348	-42.864	4956.645	226.707	393.827	8.196	867.722	0.221
29	24.651	-62.841	6173.002	272.793	556.832	9.533	820.386	0.061
30	30.622	-23.980	4864.564	387.931	447.417	11.371	911.251	0.161
31	26.648	-33.616	3860.984	218.281	602.955	8.883	1024.677	0.064
32	26.203	-24.923	4101.679	215.495	632.460	9.428	779.515	0.056
33	24.414	-42.611	5322.346	241.050	585.176	8.168	668.498	0.140

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
34	23.692	-36.625	4862.963	324.925	409.974	9.509	989.852	0.151
35	23.543	-39.187	4706.809	315.415	374.537	9.234	912.807	0.169
36	21.140	-36.687	4650.789	237.192	508.683	7.760	883.776	0.049
37	21.739	-42.134	4292.960	207.779	616.594	8.090	949.998	0.052
38	26.875	-45.103	4091.337	261.061	428.982	8.343	837.283	0.095
39	27.996	-44.945	5561.746	314.999	510.615	9.287	995.254	0.218
40	24.642	-44.488	4385.130	244.312	501.146	8.506	1233.221	0.170
41	26.991	-37.927	4135.995	264.260	498.707	10.066	884.719	0.113
43	25.165	-39.409	5881.280	319.606	537.882	9.547	896.116	0.078
44	18.798	-39.024	4332.066	248.927	464.293	8.594	870.884	0.121
45	23.948	-44.166	5134.194	302.176	491.513	10.150	1122.758	0.127
46	17.030	-46.788	4793.690	257.871	479.007	8.230	1129.144	0.199
48	29.464	-28.998	4498.665	220.623	494.960	8.629	1182.612	0.105
49	25.360	-42.421	4019.858	278.554	715.820	8.921	815.782	0.065
50	17.258	-44.935	4935.139	251.488	428.008	7.647	662.738	0.119
51	21.795	-35.351	3693.100	178.516	519.912	8.544	928.149	0.101
52	25.320	-44.782	4565.140	245.869	518.620	8.926	932.955	0.071
53	25.821	-42.875	5032.878	269.690	439.454	9.297	926.254	0.093
54	27.187	-34.786	4176.356	235.311	520.200	8.717	702.024	0.073
55	21.395	-38.757	4448.139	233.971	590.880	8.427	951.090	0.121
56	26.689	-30.174	4042.996	224.712	498.765	8.230	793.686	0.112
57	28.304	-28.324	5250.646	283.896	561.214	9.140	892.905	0.129
58	22.062	-39.233	7285.866	274.826	551.693	8.721	625.287	0.103
59	20.301	-43.145	4026.072	238.306	466.228	8.373	832.896	0.206
60	22.727	-43.829	4080.592	224.692	557.221	8.795	902.953	0.087
61	23.717	-42.219	6865.533	306.694	607.689	8.782	940.274	0.097
62	25.528	-32.729	4259.178	237.702	518.752	9.705	1010.623	0.070
63	25.747	-35.836	3771.245	226.989	572.462	9.244	939.544	0.064
64	23.566	-30.230	3907.923	202.637	535.475	8.379	928.847	0.061
65	28.262	-45.796	3753.868	309.124	520.406	9.472	935.923	0.096
66	28.842	-45.143	6335.883	290.103	449.982	10.250	931.169	0.045
67	21.840	-29.673	3961.012	233.158	472.601	8.574	1267.909	0.119

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
68	34.330	-35.409	5937.950	283.255	602.450	11.112	950.927	0.065
69	23.007	-49.658	4244.834	229.655	341.859	8.150	1130.625	0.172
70	26.437	-33.276	4191.898	250.248	502.374	9.580	1017.101	0.088
71	25.087	-33.606	5588.444	269.525	489.988	9.185	780.468	0.089
72	26.747	-28.058	5476.763	301.345	526.329	9.686	856.491	0.092
73	22.320	-46.259	4728.799	294.574	518.873	10.832	1003.106	0.098
74	25.369	-34.639	5146.504	255.706	431.411	9.238	951.544	0.117
75	20.072	-36.253	4410.922	264.005	411.942	9.405	1092.417	0.166
76	21.259	-46.499	5889.938	291.048	608.272	9.849	1032.239	0.073
77	21.303	-32.308	5134.417	290.187	430.624	8.298	1076.448	0.130
78	26.381	-42.709	4451.129	257.841	418.573	8.864	846.518	0.115
79	25.095	-30.476	3572.034	219.551	563.687	9.624	996.461	0.094
80	26.077	-22.852	4419.499	246.562	528.206	8.987	858.164	0.105
81	23.077	-44.392	4802.790	222.015	410.634	7.783	878.324	0.179
83	21.100	-30.533	3723.604	165.377	548.220	7.596	920.066	0.036
84	22.204	-44.946	4168.385	200.613	578.687	8.385	906.824	0.121
85	23.084	-26.942	7710.505	232.323	447.174	6.725	685.532	0.097
86	22.170	-36.575	4346.435	230.892	451.160	8.196	918.671	0.164
87	21.637	-45.577	4140.586	215.446	503.027	8.310	963.308	0.090
88	27.686	-35.502	5168.779	252.475	455.840	8.602	921.667	0.092
89	25.711	-55.488	5128.448	278.841	624.686	9.359	1159.863	0.089
90	28.172	-51.032	6793.194	341.559	372.516	9.844	668.029	0.166
91	35.356	-34.423	4995.167	279.850	696.140	9.768	802.044	0.072
92	22.569	-39.272	4754.340	279.394	588.445	9.527	1098.600	0.107
93	21.152	-49.608	5378.837	316.839	510.281	9.948	1158.609	0.112
94	23.031	-49.925	4943.898	269.419	701.469	9.345	1142.955	0.057
95	21.314	-56.943	5512.243	327.364	462.582	8.792	1107.557	0.234
96	22.786	-58.680	4782.017	235.090	583.039	9.931	1133.735	0.105
97	22.749	-39.846	6235.960	271.992	579.892	8.811	878.282	0.072
98	25.218	-48.180	5645.587	300.818	480.576	11.356	982.276	0.043
99	24.527	-48.904	5171.949	283.590	564.862	8.995	917.372	0.138
100	30.284	-57.233	5382.851	290.307	575.811	7.964	836.969	0.183
101	20.788	-35.043	5526.949	240.151	543.474	7.797	678.077	0.123

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
102	36.651	-37.234	7239.058	350.604	612.347	8.870	811.333	0.076
103	29.373	-29.390	4953.240	287.928	534.156	8.297	782.930	0.043
104	28.904	-47.420	5471.469	262.269	590.941	8.547	923.136	0.065
105	21.878	-57.143	5938.561	251.317	578.997	7.645	849.473	0.050
106	20.949	-49.473	6127.864	312.175	463.493	7.657	835.958	0.143
107	23.107	-51.224	5998.364	304.704	488.669	8.950	1143.865	0.112
108	20.108	-47.253	5190.875	253.744	534.149	8.382	895.013	0.143
109	29.437	-55.608	6286.161	340.978	561.628	9.471	974.006	0.097
110	30.203	-49.625	6957.850	339.696	512.819	8.908	925.533	0.173
111	25.345	-53.409	5422.910	275.344	540.404	8.863	1189.657	0.133
112	24.695	-56.407	5478.309	334.582	492.012	10.009	843.895	0.077
113	29.799	-38.584	5695.205	297.866	541.264	8.358	1134.041	0.142
114	28.779	-43.496	5207.023	317.610	559.327	10.998	974.579	0.114
115	23.251	-36.210	5616.457	265.915	528.400	7.842	1237.849	0.113
116	21.450	-50.178	5753.565	285.513	548.011	9.448	1149.418	0.082
117	27.129	-42.798	6835.297	379.274	506.258	8.887	841.981	0.067
118	25.308	-37.015	5247.059	290.406	512.433	9.971	1118.723	0.070
119	28.460	-40.601	6364.507	293.359	457.537	8.253	798.477	0.174
120	24.771	-56.033	5492.175	294.615	562.092	9.492	1126.535	0.159
121	29.297	-48.341	5120.207	290.314	598.578	8.775	812.781	0.080
122	25.970	-47.986	6262.443	352.157	645.001	9.759	1086.953	0.121
123	28.814	-49.123	6816.191	321.925	529.188	9.166	948.628	0.123
124	25.729	-57.434	7808.142	327.593	505.190	11.252	1194.316	0.098
125	21.512	-48.628	6659.430	331.332	550.433	8.649	841.595	0.142
126	21.489	-36.171	6017.700	292.800	516.135	8.830	754.484	0.087
127	21.033	-49.619	5263.550	290.921	558.493	8.083	791.344	0.059
128	30.629	-46.573	5556.291	312.760	570.723	9.752	1077.163	0.060
129	31.773	-51.449	5658.529	306.048	577.939	10.185	872.361	0.104
130	30.440	-52.132	7224.679	351.858	476.541	10.138	1009.013	0.104
131	24.616	-39.470	6906.053	286.093	536.987	7.496	907.589	0.113
132	20.912	-47.479	6569.593	285.478	405.967	8.971	808.049	0.108
133	24.930	-31.325	5216.238	244.819	677.839	9.683	895.466	0.045
134	26.933	-60.954	5723.860	312.438	578.010	7.812	938.280	0.145

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
135	22.467	-51.637	5506.660	291.855	555.751	8.764	1231.632	0.128
136	29.504	-54.570	7061.907	325.575	697.965	9.250	931.949	0.171
137	27.234	-50.106	6053.549	324.090	468.360	7.857	913.844	0.172
138	21.072	-45.922	5991.939	364.825	527.722	9.400	863.327	0.188
139	24.537	-50.111	5530.403	266.557	575.261	8.258	784.242	0.067
140	28.034	-47.051	5500.419	299.798	577.978	8.911	974.763	0.056
141	21.231	-41.777	5995.681	272.709	601.649	9.092	889.205	0.042
142	36.904	-37.198	6435.964	355.173	542.413	11.309	901.414	0.077
143	30.557	-34.567	4844.522	270.675	652.267	8.980	955.619	0.079
144	28.369	-32.836	5505.712	318.294	501.264	10.246	699.309	0.048
145	26.011	-46.916	7118.700	291.355	541.606	7.702	697.287	0.135
146	31.143	-51.182	6663.537	379.466	520.686	10.732	923.624	0.104
147	20.913	-41.841	5292.430	270.169	540.419	8.680	764.314	0.078
148	26.812	-55.568	5044.335	303.171	550.221	9.761	943.422	0.217
149	27.248	-55.280	5462.860	270.747	452.182	8.727	746.809	0.111
150	23.370	-48.466	5236.237	271.220	584.473	8.558	921.042	0.152
151	30.595	-46.406	5885.660	309.357	653.330	10.217	1052.607	0.037
152	26.841	-55.640	6805.617	297.995	581.936	10.022	829.272	0.098
153	30.649	-45.201	5724.768	287.102	555.185	8.759	786.986	0.158
154	25.744	-60.815	6044.530	286.446	559.547	7.935	992.572	0.198
155	24.159	-48.985	5584.400	273.800	544.382	8.059	1342.042	0.197
156	30.994	-47.557	7617.039	349.523	575.621	9.267	985.562	0.100
157	28.891	-46.297	5361.446	267.120	595.621	8.838	1035.932	0.079
158	27.281	-46.834	5330.952	281.785	645.387	9.274	968.912	0.053
159	30.420	-43.199	6543.160	397.690	575.122	9.849	850.352	0.122
160	31.069	-60.449	7751.540	350.119	546.182	10.343	1104.677	0.073
161	35.872	-31.914	5979.891	329.649	494.776	9.637	827.305	0.063
162	25.651	-47.816	6237.264	374.669	553.642	9.926	960.698	0.131
163	34.155	-55.083	5780.389	394.041	556.685	11.973	841.238	0.096
164	25.452	-50.847	5523.223	263.638	543.229	8.199	952.774	0.091
165	22.206	-46.938	5269.486	238.449	555.666	8.593	914.636	0.227
166	30.489	-40.386	5110.092	350.806	513.438	9.743	683.505	0.109
167	25.541	-35.761	5298.732	228.150	558.561	8.126	717.163	0.116

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
168	26.881	-43.345	6173.499	348.550	610.864	9.605	691.695	0.067
169	22.379	-42.728	5572.094	271.523	553.741	7.935	942.098	0.073
170	29.389	-51.155	7327.615	299.692	522.886	9.606	1029.764	0.094
171	25.886	-43.414	7095.084	295.991	561.556	8.104	862.413	0.206
172	29.024	-12.731	6185.681	263.709	413.351	8.500	788.478	0.128
173	27.046	0.275	4097.429	180.124	655.486	6.167	817.664	0.088
174	21.528	10.524	3729.819	161.193	553.269	5.913	756.478	0.171
175	22.592	10.981	4103.510	169.586	541.720	4.975	1010.764	0.194
176	22.830	7.306	4357.079	161.792	418.782	5.088	854.300	0.122
177	24.769	20.516	3509.649	169.711	558.685	6.348	825.461	0.074
178	22.357	24.129	4218.683	125.990	330.165	4.749	426.911	0.152
179	23.531	24.860	4011.037	153.415	490.174	4.762	757.301	0.148
180	20.311	22.561	3472.321	136.609	570.814	3.337	682.752	0.147
181	21.620	25.415	4463.514	161.570	354.932	4.001	611.399	0.185
182	22.808	15.168	4237.743	178.436	539.385	5.923	666.803	0.170
183	23.788	27.758	3408.204	146.451	437.967	5.701	917.219	0.117
184	22.297	26.502	3570.890	143.282	310.636	4.348	696.752	0.250
185	20.804	23.603	3810.668	191.407	367.326	4.223	589.624	0.146
186	22.488	22.362	3438.061	172.536	534.308	5.193	780.252	0.113
187	23.813	27.852	2794.278	169.147	397.702	6.355	436.993	0.099
188	25.985	33.189	3647.456	187.378	440.515	5.782	723.944	0.097
189	22.982	25.879	4813.585	173.439	301.060	5.144	268.302	0.158
190	19.583	30.335	3369.941	138.632	376.967	4.621	628.286	0.140
191	20.816	24.738	4303.628	210.122	390.125	5.289	778.900	0.224
192	23.709	27.402	4544.750	210.645	362.123	4.800	811.222	0.136
193	20.044	21.337	3824.289	144.389	423.587	4.407	727.858	0.152
194	20.745	30.948	5868.998	144.848	310.462	4.261	516.653	0.357
195	24.908	25.374	4142.731	193.846	527.844	6.454	816.316	0.147
196	21.621	32.434	4202.460	184.280	136.340	4.712	537.625	0.225
197	20.971	21.979	3796.718	183.541	420.836	5.045	850.180	0.184
198	21.105	23.449	3868.818	131.971	401.248	4.476	795.209	0.140
199	27.175	26.878	4211.244	227.621	396.817	7.195	847.451	0.140
200	20.577	29.767	4819.124	130.706	175.123	3.597	403.566	0.319

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
201	20.562	26.439	3342.578	141.390	335.664	5.045	723.656	0.144
202	21.806	30.931	4077.021	171.838	301.283	4.731	833.812	0.199
203	23.296	34.374	5180.416	220.072	359.126	4.869	801.637	0.187
204	20.979	23.994	2920.550	141.237	424.015	3.647	788.229	0.194
205	23.211	18.814	3726.812	146.184	382.385	4.422	751.122	0.232
206	23.176	25.964	4191.096	234.822	319.313	6.700	485.186	0.129
207	23.330	27.208	4471.880	248.728	268.478	5.955	451.442	0.146
208	24.985	36.331	4807.679	265.620	331.902	5.462	749.287	0.176
209	23.993	25.419	5861.431	185.511	354.498	6.325	607.295	0.172
211	24.873	27.274	3736.171	179.313	366.519	5.443	726.240	0.138
212	25.873	35.312	3843.620	201.089	409.665	5.155	658.507	0.186
213	23.526	32.123	3802.858	175.081	435.861	5.348	611.842	0.195
214	21.632	25.722	3327.739	133.621	355.751	6.126	656.564	0.121
215	22.900	22.121	4000.186	180.965	285.175	5.997	1020.382	0.179
216	23.141	42.610	3155.295	190.919	509.028	6.927	640.840	0.099
217	22.912	29.993	2986.521	138.437	291.933	4.752	811.815	0.148
218	25.959	26.513	3584.704	204.907	451.956	5.289	475.990	0.108
219	21.645	21.926	3860.673	167.174	246.336	4.410	1221.184	0.173
220	21.786	23.187	4614.389	163.348	258.602	5.100	821.557	0.237
221	22.820	8.555	5240.610	191.534	289.073	5.341	853.200	0.254
222	22.500	16.704	5481.021	190.124	257.656	6.057	710.488	0.207
223	22.150	9.243	4070.816	185.760	317.378	5.630	668.813	0.199
224	19.850	27.491	3705.110	182.392	354.053	4.068	775.940	0.362
225	18.473	28.251	3506.763	137.211	327.544	3.845	536.538	0.251
226	19.999	23.674	3265.590	130.933	411.963	4.565	736.784	0.203
227	23.009	25.941	4413.969	199.177	176.995	4.603	495.135	0.177
228	24.167	14.616	4484.150	156.267	211.131	5.870	523.888	0.208
229	23.594	16.525	3912.518	165.733	376.648	4.951	595.348	0.133
230	21.410	38.062	3651.022	167.426	267.166	4.049	643.706	0.175
231	22.779	27.467	6314.354	179.580	316.323	4.905	672.310	0.264
232	25.762	32.405	3540.871	192.893	473.397	5.699	618.359	0.088
233	19.771	28.986	3426.293	121.055	439.285	3.758	671.274	0.117
234	21.877	18.575	5860.654	212.955	193.425	5.367	730.218	0.154

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
235	23.620	25.263	5849.723	217.910	232.338	5.033	479.015	0.223
236	18.299	21.237	4435.383	164.647	384.928	3.303	849.647	0.207
237	22.406	23.159	4469.902	136.584	385.115	4.900	829.506	0.142
238	23.601	30.993	3242.660	173.367	203.950	6.736	693.905	0.131
239	22.305	39.029	3845.572	179.037	398.850	4.441	549.633	0.118
240	22.804	32.787	4052.282	195.532	291.863	5.765	529.482	0.185
241	22.978	35.723	3284.664	173.196	324.939	5.117	700.024	0.186
242	24.822	33.708	3757.683	289.333	433.691	5.827	1038.112	0.151
243	17.963	30.216	3628.948	150.587	222.954	3.450	1310.390	0.191
244	21.269	35.889	4370.299	180.960	402.137	3.583	720.618	0.225
245	22.771	34.470	3618.119	221.471	243.759	4.491	625.674	0.283
246	21.257	34.569	4162.157	159.931	263.121	4.309	443.947	0.342
247	22.245	39.008	3162.427	167.952	319.198	5.034	952.258	0.181
248	23.088	50.902	2698.449	164.063	305.280	6.321	864.574	0.144
249	23.829	31.053	4069.900	167.313	271.086	4.845	585.083	0.173
250	24.604	40.767	3545.879	173.498	301.912	3.754	686.064	0.152
251	24.075	36.696	2810.639	243.103	386.430	5.270	731.895	0.176
252	21.767	43.110	4390.673	212.445	355.008	4.768	543.862	0.154
253	21.014	57.372	2715.106	130.074	324.556	4.841	630.681	0.198
254	19.023	35.521	4011.067	167.015	322.514	2.723	520.103	0.159
255	22.250	43.503	4686.332	169.102	237.556	4.092	600.056	0.348
256	20.022	49.705	3910.472	203.971	290.877	3.775	967.456	0.256
257	23.247	54.567	3877.769	165.272	301.953	4.012	640.722	0.308
258	24.072	53.134	2830.304	211.217	330.452	6.109	603.725	0.233
259	22.126	71.065	3399.973	180.992	244.360	3.494	585.632	0.304
260	21.481	64.211	4160.406	236.129	304.479	3.269	718.955	0.222
261	21.157	58.040	4535.759	159.780	300.359	3.263	468.924	0.245
262	20.044	81.202	5217.704	164.803	150.397	2.137	518.340	0.400
263	25.110	60.012	2971.361	187.067	253.811	4.221	602.285	0.175
264	19.376	62.708	3424.527	178.210	357.032	3.491	634.185	0.243
265	20.230	70.830	3393.703	159.943	238.363	2.638	600.137	0.251
266	22.067	59.707	2556.200	172.759	450.410	4.242	747.748	0.183
267	21.010	71.000	2577.693	140.858	389.264	3.718	561.216	0.163

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
268	18.070	82.911	2348.533	168.313	363.848	2.176	997.185	0.316
269	17.653	73.255	3178.783	181.665	251.174	2.879	245.905	0.204
270	21.766	88.354	2697.297	171.626	235.265	2.905	740.198	0.270
271	20.589	78.560	3159.035	272.610	280.354	3.178	772.450	0.189
272	21.888	78.226	2523.012	238.455	312.950	4.440	879.645	0.211
273	19.920	82.533	2754.036	178.141	230.405	3.339	906.000	0.293
274	17.140	83.504	1830.921	148.847	165.931	2.172	666.816	0.272
275	22.224	78.299	2417.990	140.123	254.748	2.901	652.933	0.227
276	24.935	86.539	2780.558	231.093	289.910	4.142	430.881	0.190
277	20.639	87.834	2065.307	168.594	288.623	3.469	824.308	0.189
278	24.497	70.585	2236.355	152.685	304.173	4.205	522.897	0.144
279	18.628	89.629	2368.360	171.415	117.014	2.827	647.806	0.286
280	20.781	78.386	2001.878	188.925	406.860	3.556	830.530	0.243
281	21.810	77.447	2825.257	208.366	289.947	3.989	823.489	0.173
282	23.308	86.772	2391.828	184.923	442.341	3.671	910.477	0.187
283	22.498	79.942	2154.726	161.647	316.823	3.190	873.852	0.181
284	24.316	73.887	3665.527	268.887	224.842	4.461	598.305	0.242
285	21.584	69.161	2871.383	184.996	242.278	3.831	419.086	0.169
286	19.939	70.380	3125.699	168.883	261.939	3.312	816.361	0.229
287	18.163	84.930	2922.145	172.241	124.053	1.072	1010.071	0.344
288	20.261	84.603	2589.402	181.149	193.902	2.862	897.078	0.201
289	20.183	87.407	2756.766	151.039	160.421	2.445	878.811	0.259
290	24.537	70.781	2617.899	205.127	396.881	5.033	850.118	0.180
291	23.799	72.653	2633.631	173.984	336.273	4.319	831.965	0.146
292	21.606	74.364	3138.179	191.592	234.454	2.409	933.278	0.359
293	21.262	75.904	3120.949	190.087	182.330	3.089	569.685	0.293
294	20.103	76.804	2794.358	199.192	384.716	4.653	649.145	0.151
295	18.209	83.108	2238.386	152.415	209.772	2.060	605.887	0.273
296	22.258	78.771	2078.482	176.673	195.017	2.121	691.283	0.264
297	20.013	86.175	1540.566	159.135	251.071	2.575	728.938	0.202
298	22.433	75.620	3530.819	293.086	347.277	2.948	718.911	0.207
299	24.402	67.739	3339.984	251.517	260.382	7.690	696.589	0.176
300	20.412	67.405	2438.681	170.327	269.906	3.537	955.729	0.179

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
301	22.549	75.582	2628.723	224.370	169.923	3.771	909.139	0.316
302	21.755	89.393	2108.338	158.773	383.068	3.412	948.511	0.222
303	19.600	90.615	2015.238	146.241	229.746	2.313	794.231	0.228
304	21.316	73.868	3849.577	196.855	277.617	4.010	436.253	0.208
305	22.059	65.411	3068.655	163.086	260.748	4.266	793.918	0.219
306	17.384	83.561	3530.540	146.753	211.514	1.804	700.443	0.313
307	16.503	79.875	2998.494	176.893	146.608	0.897	907.588	0.190
308	22.343	69.485	3082.078	194.992	167.882	2.484	913.695	0.203
309	21.756	78.679	2730.532	211.041	247.301	4.966	679.444	0.227
310	21.750	80.890	2530.908	174.018	324.068	3.426	787.094	0.166
311	20.682	83.188	2113.268	174.838	309.885	3.831	552.746	0.140
312	19.224	76.486	3878.004	227.125	211.765	3.819	319.822	0.254
313	19.330	76.985	4590.673	227.245	215.312	2.588	662.772	0.227
314	22.533	74.126	3815.277	239.437	281.098	4.229	489.698	0.222
315	17.182	70.926	2824.635	164.200	169.071	2.448	691.761	0.347
316	20.206	85.220	2889.128	173.264	204.396	3.095	553.489	0.252
317	22.615	74.602	3567.377	216.686	193.780	3.187	569.957	0.290
318	22.884	71.317	2867.461	221.781	136.740	3.965	850.089	0.239
319	20.821	73.957	5634.231	174.076	353.602	1.635	513.991	0.263
320	18.707	79.394	2657.120	207.493	183.572	1.645	576.402	0.257
321	23.384	68.616	3236.890	281.700	215.413	4.719	714.208	0.231
322	18.621	86.764	3047.276	192.259	131.518	2.760	289.223	0.267
323	19.002	82.134	3301.181	185.055	139.077	3.667	364.446	0.290
324	22.799	77.825	3596.940	219.240	149.105	4.101	633.149	0.270
325	22.923	68.941	7965.519	282.438	285.995	4.213	804.717	0.189
326	22.948	86.870	2850.933	205.744	167.883	2.542	406.741	0.247
327	20.330	81.019	4011.993	177.538	192.983	3.884	558.127	0.294
328	20.937	75.455	3850.641	197.856	187.047	3.995	567.378	0.230
329	20.773	72.646	2580.487	173.882	165.496	3.485	715.739	0.288
330	20.867	65.035	3614.356	206.900	216.690	3.510	595.677	0.243
331	23.583	90.151	3995.754	158.062	172.470	4.180	144.767	0.234
332	23.350	86.734	3006.293	279.388	167.422	5.022	414.906	0.155
333	21.918	73.039	2193.466	187.173	487.634	4.161	639.277	0.150

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
334	21.529	75.082	2399.940	210.248	272.938	2.873	688.736	0.205
335	21.445	66.199	2799.926	218.441	216.211	3.935	767.276	0.216
336	20.441	80.828	2651.217	161.155	214.830	3.890	436.192	0.301
337	21.383	67.315	4194.895	252.602	298.491	3.954	634.244	0.284
338	17.682	70.914	2015.592	148.016	203.273	2.204	774.317	0.217
339	18.859	67.010	3864.188	209.344	239.886	3.646	628.000	0.287
340	19.450	66.143	3474.692	187.964	252.129	3.024	380.320	0.233
341	21.136	57.980	3575.755	185.656	263.663	4.898	542.838	0.297
342	20.052	70.085	2905.430	209.049	315.927	4.797	972.432	0.214
343	20.212	70.258	3164.591	184.709	234.462	3.345	953.337	0.273
344	28.199	60.899	2977.056	260.007	451.771	7.004	783.964	0.169
345	18.820	63.013	3266.967	181.459	223.491	3.506	378.208	0.332
346	25.363	55.157	3674.451	243.651	213.649	5.018	687.800	0.251
347	18.436	61.257	3200.588	192.163	215.017	3.701	930.625	0.288
348	22.724	57.619	5671.716	221.722	307.058	4.879	642.004	0.234
349	22.620	70.090	3218.970	249.918	368.442	5.603	474.555	0.254
350	26.454	59.872	3076.820	225.293	402.879	5.488	951.664	0.235
351	21.120	61.638	1699.188	140.549	337.675	4.375	871.739	0.180
352	20.351	60.018	2777.228	206.277	299.552	4.723	731.185	0.249
353	22.535	62.594	3569.085	193.888	390.470	4.832	515.160	0.220
354	28.006	53.805	5232.346	258.590	244.594	6.071	787.185	0.253
355	20.405	65.190	3332.933	216.909	268.350	4.153	883.006	0.228
356	23.005	55.060	4901.608	271.239	274.282	5.555	780.792	0.246
358	22.894	62.546	3153.264	187.540	199.308	3.156	597.785	0.257
359	22.790	67.421	3554.502	216.805	244.470	5.020	530.019	0.220
360	17.282	58.648	5352.711	169.907	226.596	2.671	182.770	0.438
361	23.681	67.508	2520.536	206.731	212.637	5.643	771.453	0.248
362	20.237	54.948	3235.387	208.803	326.548	4.828	868.600	0.295
363	18.510	61.529	2656.759	204.004	261.516	3.428	476.267	0.217
364	20.525	62.644	3285.194	199.644	337.209	3.793	490.225	0.272
365	22.436	54.962	3165.387	195.422	307.252	5.482	753.903	0.204
366	23.728	67.598	2921.980	203.506	322.779	3.766	751.674	0.180
367	19.734	60.545	3561.356	199.149	335.807	4.259	601.986	0.244

Supplementary Table 3.5 (continued).

NSFTV_ID	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
368	23.998	66.010	3612.401	221.307	372.732	4.397	710.722	0.183
369	18.475	46.630	3796.846	185.715	170.571	3.048	570.140	0.557
370	21.958	56.054	3695.250	204.918	179.627	5.305	354.535	0.278
371	22.770	55.855	3248.825	222.231	218.518	4.778	485.151	0.326
372	22.680	66.081	3128.793	215.363	201.939	4.729	467.289	0.255
373	21.700	56.757	3927.664	255.415	241.169	3.954	760.650	0.241
374	20.655	64.660	2716.543	204.009	324.178	3.342	564.601	0.248
375	20.478	67.366	3103.516	220.635	231.675	5.059	1120.974	0.314
376	23.259	62.818	2708.135	202.380	269.065	5.502	975.041	0.263
377	25.025	58.978	3793.263	232.766	384.271	5.288	879.304	0.204
378	20.104	59.365	3942.150	251.496	205.683	5.514	677.687	0.276
379	23.347	64.414	3797.454	249.509	232.617	5.163	835.080	0.268
380	23.547	61.753	3035.640	208.956	301.816	5.076	652.865	0.189
381	21.873	60.764	3200.027	163.324	337.983	4.932	478.547	0.259
383	19.623	65.831	3406.620	177.823	207.105	3.159	697.440	0.310
384	24.012	66.201	3348.150	211.670	299.244	5.993	716.313	0.249
385	19.472	56.611	3823.767	222.431	279.940	4.538	557.003	0.296
386	22.426	58.831	2925.026	215.704	273.662	6.330	955.348	0.214
387	19.478	53.663	3498.841	159.948	316.199	4.821	903.477	0.248
388	20.802	53.997	2459.226	185.518	349.970	4.478	980.167	0.206
389	26.383	66.432	3737.237	203.026	397.643	6.939	692.475	0.157
390	25.040	57.990	3299.851	249.984	431.646	7.718	816.469	0.185
391	23.545	50.571	2618.295	208.612	460.169	5.356	660.800	0.187
392	19.885	74.994	2859.267	161.285	273.249	3.529	648.936	0.254
393	21.300	61.754	4412.425	270.783	265.024	4.239	826.825	0.324
394	25.714	53.940	3855.897	219.426	343.606	5.610	665.325	0.281
395	23.063	49.092	3385.685	234.639	323.686	5.687	864.942	0.278
396	22.529	81.418	2729.374	204.103	282.597	5.574	614.391	0.232
397	20.568	57.186	3443.920	174.264	280.626	4.795	607.471	0.220
398	22.829	55.378	3902.336	242.929	367.622	5.772	699.581	0.181
399	21.488	65.742	2826.380	179.585	303.158	4.163	680.480	0.231
400	21.365	54.279	3865.677	231.283	253.462	4.919	822.351	0.297

Supplementary Table 3.5 (continued).

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
1	3.269	76.659	2488.052	3.735	0.369	2.210	22.172
2	3.060	77.243	2859.925	4.379	0.304	1.815	20.866
3	3.047	56.853	3174.446	3.654	0.407	1.746	20.406
4	3.452	62.819	4044.972	4.091	0.354	1.454	25.355
5	3.081	67.728	3008.315	3.941	0.339	1.619	25.499
6	2.798	69.820	3724.757	5.862	0.316	1.664	29.005
7	4.252	69.808	2910.324	4.109	0.423	1.510	36.443
8	2.823	59.927	3320.113	3.684	0.533	1.410	25.069
9	3.695	73.100	2597.414	6.049	0.473	1.739	21.966
10	3.495	79.655	3722.664	4.043	0.455	1.704	22.476
11	1.197	77.050	3410.106	2.700	0.465	1.334	25.103
12	2.862	84.507	3766.035	4.742	0.343	1.679	21.286
13	2.058	76.531	2445.464	10.161	0.597	1.825	26.214
14	2.341	64.710	3663.804	5.229	0.582	1.411	20.597
15	3.522	88.259	2787.914	4.761	0.351	1.595	24.637
16	3.978	91.523	3606.367	8.350	0.407	1.665	30.816
17	2.154	76.654	3253.957	3.732	0.400	1.549	20.328
18	4.087	71.798	3926.861	6.746	0.530	1.611	29.776
19	2.804	68.166	3760.301	9.871	0.448	1.504	22.726
20	2.805	72.189	2487.355	4.034	0.450	1.467	23.284
21	3.168	68.519	3598.691	5.175	0.411	1.590	22.903
22	3.522	76.168	3482.650	4.697	0.353	1.523	24.950
23	3.371	79.924	3627.365	4.514	0.413	1.530	24.086
24	4.384	75.845	2712.586	6.232	0.479	1.478	26.978
25	3.660	89.223	3459.201	4.796	0.436	1.462	27.541
26	3.566	82.733	3139.558	5.032	0.481	1.621	31.493
27	3.065	83.851	2380.625	5.885	0.553	1.561	29.665
28	2.281	89.703	3313.904	4.421	0.478	1.672	26.573
29	3.236	82.003	3327.946	3.819	0.601	1.337	31.091
30	3.286	76.355	3064.963	5.189	0.362	1.514	31.018
31	4.420	75.430	1865.514	4.856	0.292	1.747	24.029
32	4.367	82.664	2399.144	4.610	0.395	1.885	27.045
33	2.696	81.688	3370.943	5.687	0.451	1.636	28.109
34	2.758	74.965	3447.904	6.631	0.456	1.629	26.958
35	2.988	68.543	4254.165	4.174	0.599	1.599	31.951
36	3.316	81.828	3435.152	3.630	0.544	1.417	19.180
37	2.801	78.003	2350.184	5.009	0.404	1.741	24.263
38	4.393	69.053	2876.255	5.057	0.485	1.482	33.014
39	2.929	68.530	3228.572	4.410	0.399	1.640	25.865
40	2.384	72.580	3372.158	4.404	0.346	1.449	24.882
41	3.050	77.845	2034.689	5.473	0.494	1.661	36.189
43	3.039	77.367	1927.461	4.861	0.365	1.498	27.838

Supplementary Table 3.6 Summary of Bayesian adjusted shoot ionomic phenotypic values for RDP1. Per line values represent the 50th percentile of the posterior probability distribution.

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
44	3.604	80.588	3795.069	4.357	0.486	1.565	33.111
45	2.438	84.664	4155.267	4.164	0.346	1.656	26.089
46	2.042	77.772	3163.480	5.823	0.523	1.646	22.708
48	2.924	74.932	2929.049	5.191	0.382	1.595	24.927
49	3.396	74.312	3794.017	7.359	0.630	1.550	30.833
50	2.807	80.578	3737.914	5.206	0.531	1.591	27.609
51	3.582	65.607	2564.811	5.301	0.420	1.483	25.763
52	4.124	75.594	2526.485	6.165	0.465	1.573	23.068
53	2.940	72.898	3257.635	5.684	0.412	1.411	21.122
54	3.635	79.693	3933.066	6.654	0.525	1.598	23.891
55	2.473	80.119	3813.044	4.511	0.374	1.387	22.442
56	2.758	78.967	2549.817	3.762	0.410	1.601	26.336
57	2.855	68.254	2921.727	4.980	0.542	1.515	42.906
58	2.110	66.459	3422.918	4.441	0.362	1.412	21.468
59	2.345	94.065	3504.574	5.238	0.533	1.436	31.287
60	3.962	71.571	2783.057	3.707	0.443	1.508	21.486
61	1.966	73.807	3147.393	4.544	0.488	1.660	27.906
62	3.713	75.377	2505.973	3.958	0.373	1.670	26.057
63	3.497	77.028	2456.847	4.954	0.514	1.774	24.091
64	3.877	58.524	3594.388	3.607	0.462	1.340	16.040
65	3.063	71.613	3305.563	6.750	0.517	1.262	23.134
66	2.349	70.369	3672.969	4.750	0.422	1.389	26.946
67	3.513	76.402	3788.159	5.050	0.364	1.489	26.061
68	2.758	81.076	4451.517	7.230	0.369	1.642	35.626
69	2.801	79.809	3177.284	5.927	0.575	1.529	28.397
70	2.928	74.858	3060.699	6.040	0.491	1.634	34.714
71	2.526	66.946	3455.430	4.165	0.401	1.575	31.397
72	3.886	71.211	4250.369	4.933	0.451	1.285	40.746
73	2.937	72.066	2847.325	5.275	0.526	1.585	33.994
74	2.126	73.174	3412.729	3.258	0.347	1.428	25.469
75	2.871	68.235	3594.290	6.247	0.417	1.498	25.301
76	4.227	67.295	3068.717	5.188	0.436	1.525	34.901
77	1.910	64.891	3867.512	3.831	0.358	1.580	16.728
78	2.778	65.124	3227.734	5.844	0.465	1.534	30.843
79	3.491	68.140	2973.170	4.547	0.424	1.549	25.387
80	4.171	82.935	3465.029	5.139	0.551	1.548	24.727
81	2.556	82.999	2338.231	5.109	0.485	1.496	27.850
83	3.626	72.211	2239.699	4.989	0.383	1.555	24.495
84	1.357	72.095	2899.045	5.672	0.498	1.555	19.776
85	2.790	66.596	2305.027	4.831	0.493	1.657	30.180
86	1.678	64.357	2266.570	3.937	0.385	1.456	23.773
87	2.621	62.097	3678.650	4.469	0.445	1.568	28.129
88	1.865	69.202	3248.583	4.985	0.355	1.376	26.942

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
89	5.074	71.363	2977.268	5.464	0.437	1.771	39.577
90	2.379	59.236	3729.623	3.673	0.489	1.509	27.044
91	3.295	62.345	2748.886	4.334	0.465	1.227	26.880
92	4.020	78.925	3052.857	4.560	0.403	1.467	21.134
93	3.048	87.684	3798.663	5.440	0.406	1.513	22.501
94	4.154	80.329	2903.648	4.303	0.373	1.695	26.824
95	3.212	67.177	3439.203	3.667	0.507	1.337	25.139
96	3.837	73.432	2533.622	4.291	0.534	1.380	30.059
97	3.169	74.788	3374.915	6.457	0.476	1.339	24.927
98	3.293	66.528	3325.092	5.251	0.394	1.426	27.625
99	2.866	71.808	3277.534	5.201	0.589	1.420	25.389
100	3.597	82.061	2736.677	4.912	0.418	1.417	23.599
101	2.851	64.609	1944.487	3.784	0.402	1.370	20.548
102	5.123	70.549	4140.696	4.794	0.410	1.335	34.844
103	3.737	75.261	2877.990	3.937	0.342	1.637	24.987
104	2.384	75.996	3247.508	4.682	0.428	1.471	17.338
105	3.440	76.627	3182.071	5.090	0.368	1.544	29.575
106	2.095	64.182	3188.898	2.775	0.390	1.167	21.327
107	2.206	73.151	3496.718	4.883	0.483	1.370	25.859
108	2.173	75.525	3356.982	4.017	0.429	1.491	25.744
109	2.697	68.540	2388.768	5.561	0.450	1.507	30.869
110	1.987	70.858	3676.895	3.409	0.313	1.365	15.899
111	2.232	72.806	3211.440	4.668	0.461	1.544	27.356
112	2.203	73.541	2994.391	4.147	0.363	1.608	21.932
113	3.461	76.506	2084.013	4.565	0.449	1.647	27.667
114	2.327	77.551	2911.691	4.942	0.420	1.502	26.283
115	2.510	73.931	3284.945	3.536	0.348	1.386	21.522
116	3.196	71.133	3075.120	5.917	0.540	1.433	28.443
117	2.676	75.997	3544.377	3.771	0.464	1.491	26.510
118	2.396	84.016	2451.492	3.971	0.509	1.522	23.568
119	1.825	69.630	2928.433	3.022	0.444	1.340	25.780
120	3.676	72.054	3382.369	5.838	0.623	1.424	38.749
121	3.582	82.823	3525.576	3.858	0.391	1.465	22.745
122	3.073	72.086	4202.808	3.112	0.397	1.529	25.921
123	3.210	77.432	3306.124	4.764	0.501	1.447	44.378
124	2.791	82.571	3853.038	4.389	0.565	1.567	29.013
125	1.609	74.989	3489.267	3.401	0.416	1.531	20.353
126	3.078	61.911	1916.733	4.635	0.429	1.314	35.221
127	4.096	71.234	2862.631	5.401	0.443	1.377	25.787
128	3.337	72.480	2281.557	4.630	0.401	1.513	25.197
129	2.129	69.379	2482.866	5.194	0.366	1.307	29.171
130	2.310	74.236	3624.314	4.435	0.600	1.450	36.532
131	4.175	71.984	4246.627	3.861	0.556	1.561	35.839

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
132	2.394	70.566	3062.352	3.468	0.505	1.575	25.062
133	2.298	71.051	2601.415	4.240	0.500	1.386	23.830
134	1.455	82.298	3311.482	3.319	0.488	1.358	17.466
135	2.557	83.238	2757.936	5.050	0.526	1.407	27.092
136	1.769	74.901	3505.669	3.250	0.441	1.427	21.695
137	2.019	72.210	2580.589	3.664	0.627	1.264	24.360
138	2.346	65.720	3854.923	4.008	0.480	1.456	25.880
139	5.225	66.335	3523.220	5.746	0.459	1.357	21.430
140	3.555	66.107	2467.952	5.534	0.477	1.370	34.926
141	3.077	69.745	3674.530	3.482	0.417	1.504	20.626
142	2.168	68.651	3233.243	3.913	0.453	1.395	29.816
143	4.355	72.940	1930.305	4.564	0.336	1.405	27.801
144	4.099	72.536	3038.505	4.848	0.360	1.418	24.593
145	2.726	66.155	2679.142	4.160	0.429	1.396	29.401
146	2.045	68.522	3666.086	4.124	0.391	1.508	34.898
147	3.243	70.806	3205.619	5.171	0.528	1.496	25.916
148	1.817	73.865	3095.321	4.605	0.544	1.548	31.139
149	3.684	65.763	2680.144	5.077	0.505	1.454	18.736
150	2.555	67.308	2787.023	4.804	0.422	1.352	19.742
151	2.876	73.329	2702.007	3.493	0.379	1.439	25.523
152	1.795	70.251	3334.921	4.344	0.491	1.404	22.268
153	2.630	75.774	3425.668	4.229	0.515	1.344	26.333
154	1.159	60.972	3012.117	2.593	0.438	1.426	17.103
155	2.220	72.286	3153.302	3.450	0.296	1.861	21.644
156	2.095	58.753	3421.390	5.526	0.371	1.458	30.281
157	3.431	73.911	2754.489	5.597	0.412	1.349	28.669
158	3.189	79.254	2639.574	4.196	0.526	1.422	29.848
159	3.556	69.454	2980.332	4.100	0.488	1.424	26.692
160	2.232	74.050	3424.280	4.097	0.468	1.382	28.140
161	3.453	67.316	3868.599	6.388	0.466	1.531	41.736
162	3.485	71.340	3483.846	4.049	0.454	1.396	26.274
163	2.150	77.819	3258.074	4.396	0.403	1.652	21.639
164	2.621	59.042	2482.190	3.628	0.552	1.484	24.214
165	2.751	81.564	3765.026	5.233	0.449	1.552	29.458
166	3.922	72.768	3660.500	4.063	0.441	1.502	20.099
167	3.592	67.931	2084.942	5.307	0.445	1.573	18.561
168	4.309	71.150	3873.739	3.836	0.319	1.640	22.138
169	3.015	79.394	2683.816	3.399	0.451	1.426	23.104
170	3.291	71.500	3067.344	3.952	0.408	1.609	27.867
171	2.397	54.276	4027.093	2.226	0.533	1.369	28.967
172	2.078	67.310	3730.067	3.099	0.504	1.456	22.603
173	3.291	61.673	3299.891	3.466	0.295	1.628	23.843
174	3.437	62.489	3778.671	3.855	0.416	1.403	18.221

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
175	1.854	68.898	3405.358	3.011	0.408	1.484	22.059
176	3.385	62.823	3964.363	4.452	0.371	1.671	20.248
177	5.298	60.987	3875.969	3.000	0.300	1.630	13.130
178	3.388	55.111	3355.013	4.502	0.372	1.645	31.379
179	3.733	36.649	2826.450	2.916	0.368	1.221	26.961
180	2.807	44.277	3442.645	2.758	0.350	1.456	15.285
181	3.639	48.048	4228.579	2.457	0.334	1.342	22.388
182	3.831	42.630	2715.979	3.168	0.424	1.346	21.561
183	2.900	55.483	3725.614	5.018	0.433	1.397	24.895
184	3.646	56.617	4453.881	2.738	0.255	1.386	19.176
185	4.571	50.482	3923.266	5.073	0.316	1.299	22.947
186	4.002	44.105	4317.439	5.124	0.221	1.699	22.780
187	4.258	53.834	3263.959	5.536	0.474	1.343	24.942
188	2.914	58.834	3633.052	4.687	0.430	1.546	20.607
189	3.051	55.153	3571.813	3.530	0.470	1.287	22.822
190	3.775	43.412	4106.062	3.133	0.482	1.268	21.208
191	3.757	51.446	4325.555	2.839	0.353	1.443	23.625
192	3.594	67.234	3525.179	3.432	0.358	1.527	23.606
193	2.794	51.355	3953.159	3.579	0.487	1.388	22.252
194	3.208	42.791	4662.960	2.107	0.597	1.467	24.300
195	2.876	47.335	3730.376	5.230	0.522	1.532	28.387
196	3.439	41.475	3327.171	4.024	0.379	1.580	26.379
197	2.848	52.687	4455.587	3.881	0.320	1.499	21.628
198	3.055	60.039	4380.483	3.410	0.378	1.460	28.711
199	2.548	41.086	4266.591	4.819	0.241	1.370	23.254
200	2.947	55.527	4479.730	9.384	0.507	1.556	15.569
201	3.675	41.847	3541.958	3.936	0.341	1.238	27.407
202	4.109	50.525	4181.424	5.187	0.628	1.304	29.609
203	2.537	32.378	3496.757	2.772	0.561	1.348	21.971
204	3.004	47.023	3008.297	2.902	0.319	1.422	22.743
205	3.945	60.295	4695.262	3.165	0.290	1.275	24.290
206	4.068	47.477	6881.322	3.331	0.125	1.218	22.645
207	2.939	37.143	4483.643	3.155	0.264	1.411	21.023
208	4.150	39.306	4310.016	3.731	0.409	1.392	25.281
209	2.591	43.495	3108.108	3.184	0.329	1.298	19.447
211	4.635	58.651	3898.191	4.852	0.366	1.576	26.715
212	3.038	45.218	4116.885	4.653	0.410	1.374	24.437
213	4.778	51.588	4102.803	3.645	0.344	1.506	30.227
214	2.426	61.216	4196.306	4.261	0.385	1.220	19.249
215	3.906	45.942	4079.637	5.024	0.415	1.446	24.077
216	3.776	53.173	4225.449	4.056	0.192	1.369	21.049
217	3.437	45.746	2965.381	3.082	0.441	1.301	21.721
218	3.310	60.924	4567.168	4.679	0.504	1.536	28.076

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
219	3.984	35.181	3037.118	2.075	0.331	1.400	19.544
220	2.629	43.653	2933.938	2.197	0.234	1.272	18.850
221	3.538	51.925	3529.485	3.642	0.430	1.526	22.617
222	1.633	53.756	3639.493	2.708	0.355	1.264	17.475
223	4.112	59.108	4991.968	5.953	0.359	1.547	22.636
224	2.650	51.968	4271.978	2.612	0.312	1.544	19.274
225	1.833	48.447	2650.193	2.486	0.267	1.591	18.260
226	2.195	46.300	3396.836	2.748	0.369	1.645	19.836
227	2.938	42.828	3485.293	5.339	0.357	1.570	29.906
228	2.211	43.470	4138.572	2.519	0.241	1.231	26.766
229	3.436	41.491	3842.725	3.155	0.542	1.263	18.454
230	2.521	55.942	3735.158	2.927	0.397	1.513	19.149
231	2.968	35.779	4396.508	3.353	0.591	1.303	22.110
232	5.009	50.544	4296.330	4.099	0.316	1.566	25.584
233	3.971	42.333	2768.086	3.687	0.274	1.363	21.772
234	2.335	41.613	3934.130	2.324	0.381	1.240	19.801
235	3.877	44.790	3564.102	3.079	0.436	1.330	24.180
236	4.395	38.405	3261.579	3.097	0.393	1.451	29.541
237	3.254	57.075	3578.423	3.013	0.315	1.525	16.343
238	2.699	49.111	4073.404	4.564	0.344	1.569	20.562
239	3.916	37.169	3581.494	4.780	0.370	1.314	22.199
240	4.017	50.701	3889.171	3.827	0.476	1.333	22.055
241	3.482	47.787	3968.619	3.045	0.305	1.376	26.126
242	3.467	52.912	4382.703	5.419	0.312	1.454	26.756
243	2.792	44.783	3038.925	3.085	0.224	1.345	22.625
244	2.975	34.224	3240.313	2.595	0.248	1.343	21.695
245	3.091	48.851	4200.789	2.981	0.210	1.288	22.622
246	2.643	43.175	3969.084	4.096	0.405	1.501	28.006
247	3.735	51.167	3648.219	3.568	0.241	1.456	27.604
248	3.462	58.993	3607.902	3.400	0.343	1.488	21.677
249	3.083	39.536	4721.995	3.627	0.388	1.512	17.728
250	3.183	42.494	2922.170	3.560	0.334	1.503	19.672
251	3.100	56.105	3944.642	5.129	0.318	1.550	19.019
252	2.401	32.176	3554.871	2.746	0.337	1.417	16.285
253	3.756	48.783	3886.968	4.417	0.315	1.365	21.110
254	2.922	36.900	3926.754	2.049	0.183	1.338	24.382
255	2.225	54.218	3981.640	2.798	0.467	1.428	24.812
256	3.738	41.733	3289.436	2.694	0.306	1.545	18.217
257	2.679	48.858	3317.719	2.632	0.156	1.273	19.967
258	3.719	58.994	3702.749	5.795	0.431	1.479	30.371
259	2.928	52.730	5296.932	4.904	0.412	1.543	23.616
260	3.015	54.841	5019.351	2.750	0.372	1.468	25.616
261	4.300	53.220	4468.311	4.967	0.269	1.428	20.928

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
262	1.848	53.346	3898.111	4.245	0.303	1.081	22.267
263	3.485	53.984	3941.168	2.470	0.073	1.342	20.145
264	2.182	44.192	3978.790	2.801	0.402	1.370	16.623
265	3.065	53.349	3536.052	3.142	0.446	1.529	18.526
266	3.009	50.806	3440.371	3.686	0.366	1.288	29.309
267	3.603	58.441	3253.594	4.292	0.414	1.528	20.732
268	2.542	67.883	3560.582	2.732	0.333	1.505	18.263
269	3.822	52.347	3015.600	2.035	0.299	1.186	19.921
270	3.994	45.018	3463.451	2.758	0.384	1.409	24.906
271	3.361	57.632	4205.089	3.038	0.398	1.014	18.882
272	4.692	49.485	4473.454	4.861	0.218	1.224	23.418
273	4.270	55.925	3351.565	2.939	0.308	1.199	25.342
274	3.298	58.541	3836.013	5.035	0.478	1.377	17.520
275	3.297	49.600	3441.039	3.642	0.227	1.121	20.141
276	4.142	53.444	3777.703	3.515	0.315	1.382	33.684
277	4.123	55.224	3342.957	4.056	0.540	1.293	23.489
278	2.860	51.168	3215.388	4.557	0.371	1.335	23.600
279	3.337	52.040	3562.208	3.677	0.518	1.411	25.101
280	3.175	53.966	3260.463	3.307	0.281	1.508	16.735
281	2.831	70.455	3861.644	2.614	0.149	1.422	19.956
282	3.486	54.808	3798.593	4.239	0.292	1.541	22.447
283	3.481	53.761	3398.424	4.826	0.202	1.298	26.984
284	2.876	51.950	3883.912	4.046	0.257	1.294	27.868
285	3.113	49.943	3916.435	5.484	0.348	1.230	20.901
286	3.954	57.415	3655.407	4.435	0.467	1.370	26.036
287	3.188	53.221	2419.624	3.429	0.308	1.490	19.447
288	4.094	56.003	3293.991	3.157	0.319	1.356	27.436
289	3.881	65.197	3124.543	3.471	0.367	1.385	26.499
290	3.302	68.290	3847.385	4.019	0.309	1.593	19.167
291	3.284	59.122	3829.143	3.503	0.299	1.378	19.939
292	3.363	57.614	4153.584	3.212	0.419	1.196	23.781
293	2.366	55.944	3740.561	5.460	0.288	1.380	21.803
294	2.631	51.671	4851.868	4.210	0.269	1.423	19.892
295	3.238	71.159	3332.986	4.737	0.281	1.485	26.163
296	3.316	53.263	4272.943	3.094	0.292	1.360	17.609
297	4.956	62.536	3757.506	3.340	0.452	1.577	22.106
298	2.008	58.251	5046.118	2.976	0.285	1.421	16.433
299	3.070	51.625	3766.885	3.141	0.375	1.346	26.430
300	3.648	54.325	3806.728	4.436	0.392	1.427	22.229
301	4.364	55.500	3302.494	4.736	0.473	1.373	22.467
302	3.791	66.606	3912.238	3.152	0.280	1.518	17.158
303	3.786	50.668	4006.723	2.713	0.245	1.297	16.831
304	2.537	44.087	3164.743	5.198	0.411	1.188	19.793

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
305	3.415	50.826	4143.995	3.099	0.629	1.181	22.067
306	2.422	56.458	2970.881	2.171	0.358	1.298	15.031
307	3.587	51.203	3126.476	3.301	0.281	1.251	25.262
308	3.161	50.636	4279.271	4.596	0.312	1.243	27.718
309	2.231	66.018	4061.427	4.306	0.634	1.351	19.909
310	2.951	71.130	4172.612	5.816	0.363	1.443	23.819
311	3.465	69.408	3237.822	4.762	0.267	1.519	24.753
312	3.303	52.682	4258.680	3.885	0.336	1.272	21.408
313	3.416	71.556	3752.528	4.011	0.656	1.491	25.589
314	2.003	59.571	4363.791	5.260	0.468	1.337	23.355
315	2.256	44.246	4408.267	2.594	0.240	1.199	19.050
316	3.626	49.167	4382.227	5.368	0.380	1.252	28.983
317	2.610	50.909	4795.022	2.925	0.345	1.142	23.309
318	3.732	49.883	4539.051	6.876	0.422	1.242	29.805
319	2.670	59.651	4389.835	5.533	0.358	1.364	31.034
320	3.671	51.324	4588.673	4.259	0.370	1.284	21.506
321	3.894	49.309	5365.991	5.827	0.282	1.192	36.370
322	2.425	57.870	4444.581	4.975	0.402	1.388	21.185
323	2.360	56.442	4097.411	4.191	0.445	1.245	24.763
324	2.612	60.257	3932.427	5.255	0.442	1.328	27.367
325	2.317	53.706	3838.266	3.019	0.303	1.398	18.116
326	4.116	49.454	4121.753	4.694	0.515	1.235	20.381
327	3.317	57.305	3714.143	3.191	0.449	1.461	28.812
328	3.164	53.903	3922.419	4.613	0.464	1.441	22.602
329	2.575	53.575	4337.732	2.257	0.309	1.309	21.128
330	3.327	51.579	3587.728	5.612	0.523	1.328	24.014
331	2.052	58.565	4015.228	3.544	0.495	1.405	29.750
332	2.527	50.793	5502.827	4.084	0.280	1.347	23.345
333	3.410	64.988	2542.743	5.602	0.436	1.542	23.633
334	3.486	62.590	3902.841	3.379	0.376	1.413	16.976
335	4.745	57.032	3661.045	3.725	0.377	1.376	31.660
336	2.652	54.233	4144.598	4.560	0.371	1.202	28.234
337	2.206	59.928	4300.902	3.069	0.707	1.304	22.811
338	3.276	60.775	3746.259	3.563	0.375	1.254	22.480
339	2.572	57.410	3785.874	2.847	0.352	1.182	31.742
340	4.612	53.455	4117.258	3.186	0.316	1.266	23.459
341	2.306	57.042	3152.450	5.878	0.349	1.463	23.676
342	3.545	55.933	3026.403	4.926	0.350	1.340	20.326
343	3.898	61.040	2252.551	3.084	0.309	1.291	20.528
344	3.054	60.527	3252.927	4.472	0.383	1.216	19.288
345	2.507	64.647	5211.429	3.770	0.450	1.455	23.988
346	2.877	58.533	3721.892	5.681	0.338	1.406	28.602
347	5.794	66.349	2913.148	4.676	0.456	1.608	26.052

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
348	3.443	57.661	4095.346	3.243	0.617	1.336	24.652
349	2.315	59.255	3838.951	3.155	0.461	1.210	20.457
350	3.117	65.228	3566.874	6.906	0.416	1.320	28.596
351	4.424	61.203	3393.567	3.979	0.195	1.307	23.541
352	2.706	73.052	3351.102	5.142	0.385	1.214	20.485
353	2.811	65.162	3865.131	4.176	0.397	1.314	20.889
354	2.148	56.968	3402.348	6.111	0.594	1.200	26.907
355	3.622	71.579	2972.452	3.009	0.347	1.284	24.305
356	2.198	68.648	3781.667	4.216	0.440	1.250	23.328
358	2.851	56.029	3018.975	4.526	0.344	1.198	19.156
359	2.132	75.163	4507.030	5.839	0.403	1.242	24.614
360	3.236	80.198	3761.963	3.738	0.373	1.330	22.746
361	4.111	82.648	4292.146	3.231	0.426	1.296	14.895
362	3.266	81.023	3168.538	3.881	0.439	1.399	21.829
363	4.093	75.611	2831.999	4.148	0.431	1.385	22.765
364	4.474	76.971	3221.432	4.766	0.522	1.294	25.875
365	3.656	104.485	2989.494	4.177	0.401	1.393	24.286
366	2.323	66.324	3056.500	3.381	0.267	1.264	22.153
367	3.584	68.849	3995.345	3.883	0.360	1.176	33.373
368	4.779	93.598	3427.983	4.986	0.388	1.382	25.801
369	2.000	75.668	4491.658	3.816	0.764	1.299	20.233
370	1.952	73.383	3661.128	4.542	0.460	1.305	22.299
371	3.247	79.370	4041.313	4.827	0.292	1.207	20.840
372	3.306	77.915	3857.660	6.236	0.420	1.355	25.976
373	3.246	83.728	3779.774	3.750	0.317	1.243	19.496
374	3.348	68.984	2489.383	3.646	0.395	1.300	20.246
375	2.586	81.379	2915.408	4.926	0.462	1.424	20.183
376	3.286	85.736	3028.113	4.454	0.336	1.203	20.589
377	2.649	76.170	3101.811	5.496	0.526	1.287	22.746
378	2.783	73.458	4212.547	4.875	0.519	1.328	21.619
379	4.416	89.558	4606.516	4.635	0.376	1.204	23.841
380	4.225	70.146	1749.400	4.828	0.416	1.223	28.817
381	2.403	72.952	2967.881	6.529	0.429	1.338	29.095
383	2.827	74.519	2181.267	2.963	0.505	1.205	18.942
384	3.568	85.282	3860.464	6.202	0.462	1.236	25.849
385	2.848	96.397	3612.965	3.045	0.516	1.243	18.721
386	3.373	80.541	3257.650	4.577	0.421	1.447	21.589
387	2.381	75.495	2111.106	4.520	0.419	1.322	28.959
388	3.824	75.281	3161.934	4.649	0.390	1.209	27.483
389	3.510	78.077	3035.105	4.099	0.425	1.318	28.143
390	3.003	74.829	2550.673	5.974	0.438	1.318	31.167
391	4.062	72.163	3119.241	5.902	0.460	1.304	26.099
392	2.564	71.415	3456.724	4.887	0.520	1.259	21.076

Supplementary Table 3.6 (continued)

NSFTV_ID	As	B	Ca	Cd	Co	Cr	Cu
393	2.597	68.909	2914.643	3.772	0.402	1.210	24.568
394	3.792	90.972	3190.533	4.852	0.432	1.143	23.230
395	2.099	77.898	3410.644	4.556	0.648	1.301	19.656
396	2.900	81.115	3253.800	4.203	0.370	1.320	20.865
397	3.072	75.048	2518.914	4.597	0.406	1.327	31.164
398	3.865	73.727	3186.972	4.708	0.449	1.268	34.515
399	3.347	76.742	3575.923	5.278	0.516	1.341	21.380
400	2.579	76.479	3500.361	3.205	0.452	1.249	23.440

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
1	105.524	0.971	44552.714	0.793	4358.382	52.347	5.006	301.693
2	123.314	1.230	45686.057	1.050	3180.950	49.985	3.936	324.565
3	80.208	0.830	47190.264	0.552	3615.835	38.399	1.382	225.303
4	148.859	0.757	48231.303	0.916	4448.342	73.996	3.443	263.743
5	120.446	1.929	43644.924	0.853	4118.097	77.507	4.566	195.451
6	141.463	0.595	47796.140	0.785	4030.854	70.293	4.114	276.965
7	159.607	1.158	45105.675	0.959	4220.098	64.643	11.445	343.458
8	118.988	1.439	47665.510	0.845	5327.009	57.043	5.492	219.165
9	145.020	1.291	50019.578	0.563	3606.394	62.276	5.462	240.809
10	130.607	0.672	47078.827	0.783	4276.284	80.535	3.756	256.087
11	91.926	0.790	44920.396	0.757	3089.547	28.433	3.341	283.304
12	84.725	1.689	44822.812	0.896	4483.366	59.090	3.998	162.813
13	123.890	0.737	52068.942	0.853	3804.649	50.817	3.087	224.805
14	159.697	0.869	45864.417	0.827	4180.151	61.028	3.137	314.182
15	122.051	0.877	44857.030	0.808	4067.114	62.298	4.122	215.363
16	140.134	2.350	45906.783	0.921	4740.579	79.048	4.125	257.508
17	106.006	0.809	41705.820	0.481	3489.349	53.845	3.468	268.707
18	129.722	0.979	46183.129	0.719	4242.623	72.601	4.680	248.523
19	135.471	0.784	50455.153	0.759	4731.988	79.784	5.979	168.584
20	143.366	0.066	46169.275	0.819	4468.397	86.177	3.133	291.512
21	112.520	0.599	45221.743	0.484	3573.214	60.213	2.631	238.286
22	123.104	0.029	46023.334	0.814	4049.712	56.464	4.387	275.182
23	121.691	0.865	44452.237	0.701	4958.907	66.723	5.222	259.196
24	161.154	0.085	45586.358	0.904	3477.787	62.727	2.523	274.409
25	172.695	-0.078	44397.054	0.992	4626.377	82.070	3.870	277.971
26	137.223	0.327	48179.537	0.864	4868.431	56.788	3.140	270.076
27	174.069	0.913	47048.832	0.943	4598.336	78.810	2.315	332.079
28	100.961	-0.124	45652.524	0.656	3465.837	66.373	4.557	153.445
29	93.660	0.208	44310.126	0.985	3407.786	65.500	1.304	236.917
30	130.173	0.725	45811.166	0.728	4399.669	60.346	3.064	300.536
31	96.455	0.324	48183.134	0.974	2981.781	63.141	2.961	320.603
32	141.018	0.524	48799.730	0.785	3857.365	64.839	3.796	304.115
33	132.463	0.393	46443.907	0.862	3677.793	93.766	3.931	250.338
34	96.233	0.533	45198.662	0.764	3565.145	49.297	3.046	275.335
35	95.489	0.187	45697.937	0.666	4414.839	58.261	3.773	225.075
36	150.045	0.501	45702.735	0.775	4559.259	59.721	5.871	267.612
37	129.836	0.354	44913.844	0.956	3782.862	73.687	3.813	274.531
38	155.047	0.577	44738.722	0.761	3638.859	62.251	7.143	332.580
39	88.532	0.761	45872.109	0.640	3910.610	51.941	3.705	326.227
40	140.529	0.105	45796.679	0.908	4360.915	76.809	5.781	267.676
41	135.486	0.108	46579.730	0.917	4159.786	70.787	7.030	284.130
43	136.773	0.154	35561.522	0.762	2571.708	66.957	4.132	238.028
44	137.918	0.381	47227.840	0.772	3782.266	57.923	3.903	189.814

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
45	130.834	1.115	49463.963	0.962	3555.596	78.996	3.548	194.271
46	124.994	0.291	49273.626	0.827	4740.017	65.580	2.533	302.879
48	125.278	0.617	45385.230	1.191	4163.462	67.113	5.045	270.377
49	151.644	0.245	48222.487	1.168	4071.229	87.668	4.327	328.447
50	136.063	0.935	49060.950	0.733	4102.785	71.755	6.525	157.433
51	114.795	-0.204	46385.741	0.752	3711.048	73.735	4.021	279.658
52	131.749	0.499	45630.511	0.887	4020.553	62.555	3.312	315.998
53	116.127	1.445	46531.133	0.902	4519.279	83.526	4.352	206.018
54	180.024	-0.036	45046.281	0.804	4849.291	85.066	3.405	301.403
55	127.367	0.123	46746.472	0.778	4647.546	62.550	7.683	214.936
56	90.187	0.653	47181.107	0.944	3714.166	63.055	2.976	250.017
57	102.160	0.679	45746.262	0.589	3999.218	56.797	3.537	258.275
58	116.430	-0.284	49626.293	1.024	3998.278	62.321	3.712	278.150
59	177.107	0.551	51299.896	0.831	4422.897	81.393	5.119	257.190
60	157.354	0.062	42660.596	0.721	3818.378	74.986	4.194	288.861
61	104.602	0.362	41183.339	0.840	3349.763	60.710	2.881	281.448
62	133.330	0.412	47226.797	0.796	3812.743	82.870	3.541	296.735
63	141.229	0.497	45786.906	0.811	3904.053	70.567	3.492	354.387
64	161.646	0.006	45860.919	0.741	4339.470	80.413	5.026	304.375
65	165.186	-0.263	47028.846	0.809	4563.869	74.911	3.039	251.506
66	103.758	0.801	46485.788	0.705	3599.792	53.812	2.736	278.788
67	130.578	0.542	54679.950	0.754	4767.860	92.581	4.088	294.049
68	180.413	0.781	46049.023	1.042	3558.909	120.906	4.071	358.772
69	184.043	0.055	45911.176	0.842	4394.291	84.804	2.741	259.011
70	129.902	0.815	45789.035	0.893	4289.723	64.157	9.551	271.238
71	76.381	0.541	43641.670	0.677	3500.046	39.820	3.987	338.994
72	145.970	0.398	47542.223	0.748	4973.372	64.292	3.350	343.265
73	153.001	0.324	48689.323	0.954	5010.605	57.108	3.196	259.661
74	127.469	0.179	44030.877	0.865	4078.708	53.675	3.132	280.666
75	116.983	0.815	46638.894	0.849	3619.754	64.540	5.212	243.810
76	121.600	0.174	44042.320	0.752	3596.539	61.254	3.118	313.855
77	80.752	-0.369	49425.495	0.767	4543.440	46.180	2.181	269.365
78	138.191	0.488	49906.651	0.876	3959.132	80.000	3.770	223.845
79	124.096	-0.246	47522.020	0.838	4271.870	63.232	4.192	295.524
80	172.695	0.313	43818.518	0.839	4295.015	76.097	3.804	396.405
81	103.488	0.624	49730.712	0.878	3530.286	77.093	4.499	197.366
83	97.333	0.114	48219.886	0.723	3729.417	59.580	2.626	312.713
84	117.568	-0.357	46952.404	0.883	4543.565	64.173	3.642	265.089
85	129.025	0.118	46984.148	0.821	2999.945	68.229	2.905	156.941
86	125.244	0.398	47293.612	0.816	2894.469	71.688	3.922	209.923
87	130.255	0.477	45388.913	0.820	3866.790	85.471	4.940	257.554
88	153.836	0.381	48760.720	0.843	3800.441	85.615	3.619	238.486
89	144.558	0.924	46364.099	0.907	4599.237	70.194	3.958	306.864

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
90	106.891	-0.099	53029.020	0.669	3014.759	41.855	2.270	264.620
91	142.990	0.530	42465.644	0.829	3596.056	79.485	4.104	287.007
92	182.191	-0.017	49417.086	0.788	3813.834	76.103	3.260	271.471
93	119.812	1.023	45457.778	0.969	3739.171	71.637	3.343	204.402
94	121.857	0.035	48432.254	0.868	3890.447	69.838	3.718	337.107
95	111.434	0.177	44103.377	0.745	3542.303	47.613	9.385	227.382
96	178.644	-0.021	45342.241	0.925	3949.944	104.388	3.502	309.022
97	138.553	0.151	44735.150	0.713	3222.625	55.892	2.095	291.722
98	148.318	-0.200	54454.051	1.006	5782.385	71.507	5.178	328.595
99	154.157	-0.114	49594.398	0.618	4044.220	55.785	3.126	248.811
100	157.344	-0.263	45541.718	0.702	3589.704	73.358	4.557	301.786
101	115.464	-0.289	46819.739	0.922	3147.103	43.239	3.284	331.237
102	113.059	0.304	37081.691	0.891	4158.241	61.459	4.341	311.999
103	114.520	0.172	46269.496	0.939	3804.403	76.601	4.240	299.530
104	145.023	-0.331	46358.286	0.969	3699.903	78.278	4.345	299.282
105	168.935	-0.224	45595.488	0.982	3533.987	89.880	3.976	332.850
106	94.807	0.088	45407.637	0.655	2193.311	39.244	3.496	166.718
107	167.654	-0.010	46583.356	0.851	5010.878	90.192	3.155	278.499
108	112.732	0.265	46712.110	0.848	4537.785	56.157	4.724	307.495
109	136.838	0.098	44950.656	0.940	2949.641	42.386	1.480	296.011
110	91.062	-0.169	43839.417	0.717	3516.741	42.578	3.945	233.457
111	153.265	-0.403	46300.287	0.843	4443.580	94.204	2.841	281.016
112	128.521	0.290	46744.458	1.166	4068.566	64.885	2.548	227.442
113	145.831	-0.148	47563.311	1.057	3144.691	82.796	4.192	231.371
114	157.475	-0.037	45870.417	0.700	4084.985	72.296	2.837	230.237
115	113.362	0.319	45957.782	0.521	3892.627	62.050	6.176	193.707
116	237.298	0.149	47725.064	1.056	3972.621	134.672	2.958	292.579
117	118.668	-0.166	43492.469	0.717	3079.782	60.250	2.347	305.056
118	143.578	0.225	45197.639	1.033	3352.162	71.067	5.473	255.798
119	98.972	0.018	44601.524	0.546	3111.545	39.475	2.722	250.481
120	181.926	0.340	44745.451	0.846	4094.204	67.121	3.677	268.908
121	115.057	0.170	46784.053	1.038	4069.448	65.725	3.570	335.319
122	123.498	0.717	44247.450	1.053	3962.213	65.099	7.611	268.148
123	125.480	0.745	41882.685	0.789	3775.667	66.242	3.639	331.321
124	189.410	1.083	44377.545	0.962	4032.954	85.010	4.057	291.113
125	99.789	-0.029	44540.376	0.614	3531.043	45.001	2.728	246.802
126	101.953	0.309	44235.031	0.490	2428.508	49.319	2.745	327.280
127	130.890	0.276	46264.327	0.850	4240.833	69.920	4.219	259.849
128	156.740	0.056	46486.216	1.007	3514.840	79.905	4.493	304.516
129	112.618	0.404	41623.055	0.897	3235.786	62.585	3.791	354.884
130	110.561	0.653	44284.423	0.770	3368.208	63.584	3.768	226.421
131	99.882	0.511	47033.471	0.902	4030.410	58.993	3.458	464.133
132	77.664	-0.042	28175.531	0.894	2599.211	47.852	2.811	216.158

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
133	140.302	-0.251	47436.173	0.940	3292.741	68.542	4.174	228.902
134	128.447	-0.196	46938.302	0.852	5235.680	52.799	2.472	219.215
135	161.948	0.058	45716.102	1.011	3359.356	89.911	2.632	300.772
136	87.228	0.407	44996.526	0.689	2656.744	41.340	3.328	213.385
137	83.843	-0.152	44481.889	0.735	2716.789	43.121	3.209	204.801
138	100.467	0.329	47688.062	0.568	3623.819	37.497	3.940	249.310
139	177.330	0.300	47361.516	1.036	4768.555	65.554	3.543	319.614
140	198.503	0.145	43825.367	0.741	2893.653	73.622	2.947	315.946
141	112.033	0.407	42883.929	0.760	3478.940	47.756	1.837	406.416
142	97.930	0.375	44382.200	0.709	3696.263	60.794	4.132	206.342
143	119.713	-0.132	45604.236	0.669	3134.348	66.254	4.574	316.800
144	135.136	0.022	46442.116	1.046	3811.439	82.601	4.927	281.352
145	95.356	0.405	44760.179	0.690	3090.775	69.839	3.565	263.908
146	91.348	0.810	43762.177	0.790	3624.622	54.255	3.997	217.410
147	174.384	0.181	45451.642	0.843	4171.778	71.716	4.788	320.380
148	109.174	0.000	46121.486	0.815	3151.917	55.026	3.252	240.224
149	182.285	0.430	44423.046	0.896	3640.358	61.675	5.866	237.099
150	123.339	-0.170	47406.163	0.804	4238.852	58.155	3.207	233.216
151	124.310	0.133	45547.742	0.919	3175.119	70.445	4.898	276.846
152	118.600	0.237	47797.194	0.964	2885.044	71.380	5.396	181.808
153	147.420	0.483	46826.074	0.766	4000.677	66.406	4.779	243.827
154	105.044	0.030	47900.558	0.846	6027.747	57.890	4.901	245.029
155	118.263	0.154	54069.671	0.910	4534.716	60.384	4.486	248.062
156	128.087	0.656	43868.946	0.858	3387.799	59.531	3.509	344.500
157	141.916	0.412	47085.289	0.980	3678.754	77.805	4.823	283.169
158	135.149	0.507	48000.518	0.939	2967.362	63.380	4.088	299.883
159	99.299	0.192	44431.627	0.985	2912.822	47.508	2.972	307.151
160	169.718	2.253	45020.525	0.892	3573.476	71.683	3.673	220.903
161	157.701	0.735	42886.069	0.812	3683.780	92.762	3.355	325.694
162	110.832	0.568	42963.753	0.683	3975.376	52.770	4.151	357.270
163	112.516	0.145	43902.941	0.668	2863.174	49.896	2.974	278.146
164	145.415	0.601	46808.010	0.758	3563.752	55.996	5.039	386.931
165	123.753	1.187	46031.778	0.870	4076.710	57.314	5.735	275.517
166	140.458	0.093	43672.282	0.837	4079.690	61.190	4.077	290.336
167	114.157	0.154	46802.345	0.858	3132.312	49.449	3.506	322.254
168	154.727	0.164	45041.229	0.821	3954.765	72.101	5.351	295.913
169	145.696	0.131	45837.810	0.926	4034.877	70.654	5.394	224.126
170	119.432	-0.068	45733.950	0.970	3325.388	60.483	4.074	268.195
171	141.067	0.509	44170.085	0.646	3401.534	51.628	3.422	209.555
172	118.760	0.265	45914.543	0.737	3161.503	59.681	4.613	196.072
173	144.393	-0.036	53823.981	0.939	3693.894	71.938	4.735	298.096
174	149.433	0.498	44948.198	0.825	4907.451	39.063	5.329	267.831
175	120.656	-0.159	47270.505	0.858	5170.174	50.826	3.566	193.360

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
176	202.009	0.346	40549.858	0.802	4475.178	73.617	4.898	284.086
177	129.213	0.737	46875.441	0.809	4083.715	57.803	4.637	322.995
178	177.214	0.996	48327.861	0.767	3082.046	72.167	5.016	376.419
179	134.223	0.840	44939.152	0.786	3927.079	38.119	5.270	250.312
180	165.274	1.049	34793.142	0.743	3912.961	69.869	5.225	262.958
181	173.629	1.209	42869.900	0.553	4059.722	38.411	7.106	209.685
182	155.946	1.198	41468.947	0.636	3491.714	36.003	3.121	337.750
183	187.507	1.100	50150.389	0.776	4511.173	73.437	3.426	309.231
184	166.236	1.958	47849.789	0.480	4512.639	55.654	6.279	252.960
185	195.472	0.814	43161.219	0.635	4453.349	65.527	5.043	341.863
186	221.899	0.381	43786.018	0.729	4186.200	99.870	6.888	338.929
187	216.287	0.649	44956.163	0.717	5089.599	75.245	5.865	313.872
188	221.059	0.953	39055.418	0.769	4990.266	72.498	6.125	291.755
189	164.734	1.016	46525.394	0.635	3887.108	65.071	3.930	341.798
190	158.620	0.482	41714.030	0.838	4742.682	71.504	6.244	293.402
191	194.201	1.682	39751.556	0.592	4482.486	75.570	4.708	288.977
192	195.700	0.840	55188.511	0.709	4401.256	70.086	5.155	296.210
193	163.431	0.812	42896.800	0.635	4415.241	68.655	5.672	300.628
194	140.430	0.737	38201.236	0.555	5031.553	56.711	7.550	243.932
195	236.948	0.812	41886.228	0.689	5248.756	70.281	4.221	273.087
196	149.817	1.039	38828.689	0.429	3418.170	52.364	4.118	356.252
197	182.187	0.848	40685.669	0.696	4911.285	72.432	7.119	321.801
198	175.495	0.138	46702.296	0.703	5132.027	92.582	6.644	305.455
199	154.290	1.038	45179.397	0.749	5716.508	79.118	4.714	268.140
200	153.079	1.308	49467.345	0.531	5079.359	51.960	6.074	256.276
201	166.185	0.937	44204.062	0.741	5019.208	60.173	6.562	322.467
202	221.941	0.704	44339.468	0.595	5811.630	96.282	5.070	287.705
203	127.033	1.065	42201.628	0.559	3466.137	35.806	4.600	322.100
204	154.192	0.481	44957.378	0.779	4330.987	44.340	4.546	309.382
205	193.804	1.437	40904.699	0.408	5803.964	73.255	7.225	325.562
206	143.027	1.688	43397.262	0.582	5289.136	72.626	7.267	255.782
207	137.793	0.822	40172.328	0.564	4046.128	47.620	6.099	315.541
208	154.757	1.333	43378.974	0.513	3685.403	45.665	4.954	283.344
209	116.084	0.625	45247.837	0.700	3506.534	46.583	3.897	297.463
211	292.146	0.793	39155.823	0.729	3970.168	117.823	6.071	379.913
212	221.161	0.450	40496.375	0.653	4862.616	86.955	4.768	267.517
213	163.804	0.829	48155.088	0.676	5900.944	69.126	4.461	249.364
214	184.786	1.067	47361.025	0.922	4673.614	60.469	4.289	505.687
215	225.772	1.289	46305.648	0.740	5530.019	89.038	4.061	408.526
216	167.050	0.979	45953.343	0.738	4459.726	108.667	5.956	334.083
217	179.289	0.733	43238.641	0.569	4663.567	64.246	6.788	293.078
218	220.174	0.695	55472.215	0.819	5628.822	77.689	7.122	250.670
219	147.546	1.408	47956.053	0.575	5659.233	43.173	5.773	283.411

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
220	128.561	0.964	47003.169	0.608	3350.413	49.156	5.582	255.297
221	180.550	1.738	42384.285	0.652	4432.239	60.498	4.661	256.291
222	107.110	1.214	43130.790	0.618	4561.342	46.503	3.854	241.529
223	199.545	1.412	46400.661	0.755	6326.240	66.418	4.256	367.792
224	159.127	0.929	43732.451	0.733	4918.952	78.384	8.303	208.729
225	131.097	0.177	47866.080	0.636	4001.454	45.575	5.445	227.545
226	165.275	0.156	43331.509	0.606	4061.404	70.109	3.404	331.618
227	185.141	0.419	45636.181	0.567	4050.873	66.109	4.438	313.680
228	133.558	1.253	50756.840	0.807	4734.089	53.035	6.836	206.320
229	154.710	0.976	44954.270	0.666	4810.675	65.604	6.446	280.499
230	164.497	0.621	47504.524	0.631	5137.490	57.572	6.169	249.592
231	134.275	1.234	46884.121	0.494	4157.469	46.172	4.581	247.221
232	175.800	1.031	48961.957	0.938	3760.164	93.505	5.028	334.508
233	150.112	0.525	43566.902	0.682	3905.939	47.595	4.061	300.529
234	140.755	1.076	45458.292	0.527	3825.237	42.690	2.893	236.653
235	136.461	1.133	45646.943	0.569	3456.625	39.758	4.754	326.513
236	145.619	0.584	39762.147	0.665	4304.018	67.834	4.055	299.505
237	163.836	0.652	51030.422	0.608	4909.702	67.753	4.790	465.461
238	219.961	0.895	45775.263	0.793	5004.673	97.285	4.004	285.317
239	221.826	0.784	47163.302	0.681	4734.697	69.155	4.066	300.218
240	197.256	0.764	46095.734	0.542	4806.406	64.330	3.897	312.032
241	128.135	1.390	46814.904	0.600	4227.713	49.160	4.859	259.811
242	262.338	1.192	53942.984	0.725	5340.562	88.882	5.440	265.517
243	170.198	1.366	45154.979	0.593	4298.335	64.365	6.102	267.883
244	166.488	1.335	46701.572	0.566	3849.910	63.201	5.915	247.683
245	178.182	1.496	45536.787	0.578	4651.186	50.354	6.188	282.233
246	164.859	0.721	42782.179	0.734	4093.689	63.390	6.683	223.283
247	203.596	1.122	43749.219	0.758	4577.566	73.985	6.136	270.902
248	159.429	0.889	51200.635	0.842	4512.754	66.577	5.845	310.895
249	200.878	1.397	47574.937	0.382	4228.077	89.352	8.242	284.649
250	166.517	0.601	49429.871	0.746	4056.074	63.270	5.021	254.498
251	194.241	0.555	47090.563	0.832	6125.129	68.111	3.896	293.051
252	163.004	0.503	45403.134	0.395	4443.800	48.484	3.763	266.657
253	184.317	1.094	51527.162	0.701	5204.842	78.442	3.868	299.275
254	134.108	0.983	45492.603	0.495	4607.643	63.851	2.878	240.671
255	128.497	1.157	45174.181	0.535	4086.354	33.840	5.109	259.559
256	169.251	1.188	48270.093	0.687	4695.158	69.875	6.567	300.181
257	136.835	0.931	44011.582	0.454	4075.436	59.597	5.400	216.318
258	200.776	1.711	46087.898	0.511	5089.473	67.231	5.268	344.859
259	176.908	1.628	54484.857	0.672	4715.686	94.390	6.170	205.336
260	173.286	2.934	43867.896	0.644	5014.053	77.240	5.657	192.073
261	187.695	1.734	48455.478	0.600	4085.046	49.413	5.540	427.246
262	159.626	1.517	48628.195	0.530	4309.245	61.006	4.208	257.892

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
263	147.478	1.411	44730.237	0.674	5131.635	64.055	7.639	244.337
264	163.838	1.747	48366.458	0.512	4868.146	54.509	5.166	277.742
265	155.440	1.674	49270.714	0.628	4397.876	60.444	6.008	365.040
266	196.281	1.740	47377.110	0.707	4122.692	79.758	5.119	322.554
267	153.905	1.180	46648.201	0.894	4157.347	84.071	4.349	347.636
268	153.974	0.592	48693.202	0.881	5048.703	65.312	5.537	274.424
269	106.737	0.944	49536.890	0.479	3893.773	40.332	4.638	531.178
270	178.602	0.543	45444.011	0.714	4353.609	68.133	5.270	376.219
271	159.653	1.444	45482.835	0.656	4936.095	62.904	5.805	339.466
272	207.809	0.859	44481.606	0.620	5820.419	94.167	3.782	302.791
273	168.701	1.095	47361.357	0.732	4621.644	75.050	5.763	232.645
274	177.098	0.850	47051.919	0.582	5505.400	61.273	5.039	250.065
275	133.365	0.962	46792.294	0.645	4215.428	56.078	4.758	340.670
276	178.676	1.019	46582.862	0.977	4020.451	88.137	7.624	449.236
277	159.512	1.073	47112.457	0.809	4629.794	89.154	5.676	312.153
278	139.726	0.645	45669.133	0.654	4310.968	77.989	3.721	327.823
279	157.295	0.651	46676.079	0.682	5122.004	50.687	4.567	336.705
280	217.085	0.866	46064.355	0.859	4491.984	78.078	4.744	265.206
281	143.362	0.805	42369.980	0.842	5086.870	62.380	6.105	216.812
282	202.609	0.482	48293.321	0.895	4874.095	74.231	4.955	365.177
283	163.507	0.765	46414.032	0.681	4633.126	67.896	3.751	418.142
284	165.661	1.069	43999.686	0.533	4191.887	52.182	4.672	232.240
285	166.878	1.031	46280.044	0.700	5238.733	80.277	3.805	401.240
286	180.644	0.942	46400.068	0.775	5198.073	63.521	4.168	323.698
287	137.314	0.672	46618.913	0.781	3758.784	58.461	4.737	246.062
288	184.575	0.921	41635.317	0.774	4969.104	74.683	6.498	344.137
289	162.204	0.829	45202.196	0.734	4461.861	74.102	5.263	311.994
290	162.896	0.324	46097.814	0.901	4211.024	75.641	4.498	331.499
291	177.688	1.255	46869.158	0.625	4089.096	85.846	5.295	264.620
292	177.503	1.072	47197.473	0.568	5413.320	56.250	5.128	242.050
293	194.523	0.898	46901.252	0.418	4239.738	68.605	4.307	310.391
294	157.806	0.766	42636.090	0.779	4607.035	69.431	5.080	231.521
295	171.778	0.391	46060.773	0.722	4911.974	57.972	6.544	335.221
296	188.324	0.902	44909.499	0.632	5477.576	85.110	6.661	256.978
297	151.788	0.903	47795.447	0.693	4606.042	64.426	5.818	359.694
298	159.445	0.687	52534.752	0.711	3959.169	61.634	6.153	302.318
299	129.225	0.448	44144.730	0.951	4153.714	50.721	2.409	360.686
300	173.292	0.858	42667.860	0.705	4373.570	55.751	4.773	326.975
301	177.433	1.064	46891.602	0.687	4367.240	69.068	6.625	237.480
302	173.948	1.185	52761.842	0.598	4350.271	48.313	5.167	303.882
303	169.646	0.814	45473.396	0.512	4927.970	75.028	6.833	258.607
304	143.763	0.814	46727.413	0.571	3381.679	52.705	2.843	226.984
305	173.060	1.208	45536.839	0.681	5072.391	66.631	5.780	385.750

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
306	134.860	0.958	46776.112	0.754	4539.044	59.082	5.632	286.137
307	181.449	1.149	46163.995	0.713	4470.408	66.086	5.865	316.269
308	194.084	1.207	45551.065	0.893	4472.888	72.932	4.483	288.099
309	188.463	0.634	44392.245	0.869	4672.372	65.159	4.747	283.562
310	203.064	0.628	49356.263	1.048	5721.901	135.847	4.010	325.632
311	142.824	0.649	48044.643	0.701	4028.453	58.417	4.529	293.871
312	169.514	0.902	50023.754	0.696	3619.776	55.938	5.173	238.498
313	155.617	1.262	43185.270	0.670	4129.296	55.058	4.906	359.387
314	189.724	0.883	48762.167	0.724	4423.793	69.783	5.630	284.846
315	123.649	0.991	47268.707	0.573	4597.524	37.803	3.717	222.659
316	182.708	1.003	48667.766	0.823	4260.232	78.909	5.952	238.192
317	177.722	0.985	48956.140	0.644	4436.613	84.896	6.903	200.991
318	228.093	0.971	48641.284	0.508	3815.303	93.107	5.319	239.097
319	188.601	0.916	49027.614	0.788	4456.417	98.524	7.263	219.239
320	171.639	1.060	47111.223	0.571	4236.402	71.963	5.734	254.334
321	204.409	1.349	47631.255	0.673	4292.715	88.759	7.734	258.974
322	162.384	1.013	49057.453	0.542	4664.763	75.477	7.169	232.148
323	149.345	0.915	49496.226	0.498	4464.280	49.774	5.309	267.238
324	172.155	1.074	47877.543	0.706	3530.761	79.850	7.972	198.026
325	136.588	1.413	45734.942	0.526	4081.308	47.126	5.267	263.931
326	204.306	0.999	48069.354	0.677	4003.991	66.597	7.264	333.930
327	166.206	1.037	48584.213	0.642	4127.393	63.299	5.646	224.019
328	172.774	0.979	47483.333	0.679	3799.559	77.378	5.422	321.582
329	131.455	0.974	47351.706	0.586	3749.917	61.735	6.692	185.038
330	160.979	0.995	48888.925	0.745	4511.201	82.989	4.241	314.702
331	207.285	1.111	48217.455	0.654	3896.224	82.882	7.035	231.911
332	185.361	0.981	44119.517	0.452	4968.973	80.241	5.996	301.538
333	158.405	0.636	47629.212	0.933	3675.220	71.566	4.185	306.982
334	177.745	0.837	47409.909	0.859	5149.426	63.864	5.262	291.505
335	264.452	1.045	44185.186	0.758	4642.156	96.754	5.357	281.473
336	145.812	0.581	49477.266	0.688	5078.065	68.911	6.566	254.277
337	128.491	0.547	51931.084	0.696	3026.146	42.885	2.914	246.216
338	186.254	0.956	48140.971	0.693	4410.511	62.900	5.069	257.062
339	140.022	0.787	46049.802	0.748	4109.934	49.122	2.836	270.784
340	158.508	1.032	46603.977	0.722	4999.335	55.846	4.461	279.135
341	183.830	0.918	50268.011	0.979	3580.751	65.514	4.649	289.196
342	183.578	0.758	46700.012	0.759	4828.106	77.418	3.801	281.448
343	137.359	0.508	47104.064	0.817	4203.132	61.556	3.757	304.560
344	210.150	0.794	46792.532	0.615	4231.272	65.189	4.811	341.247
345	172.258	0.577	57269.153	0.597	4180.923	74.090	6.552	172.190
346	178.490	0.981	50680.989	0.659	4128.855	79.859	5.157	269.287
347	151.334	0.661	49238.938	1.124	4404.251	66.417	4.183	372.639
348	138.810	1.045	45691.910	0.563	3601.581	43.544	3.007	255.843

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
349	128.057	0.809	46015.227	0.695	3229.409	39.145	4.188	226.281
350	235.592	0.598	45481.275	0.784	4367.033	102.059	6.690	308.397
351	140.716	0.696	43092.578	0.750	3474.998	67.696	5.311	310.119
352	193.366	0.957	52429.263	0.756	4364.014	74.612	4.106	284.302
353	176.796	0.649	48666.181	0.616	3278.475	72.588	5.405	259.094
354	152.567	1.205	35638.356	0.587	2704.643	57.609	4.330	228.836
355	141.906	1.201	48489.245	0.743	4521.294	40.626	6.290	259.624
356	149.890	1.224	44207.235	0.894	3220.447	52.484	3.722	244.964
358	188.801	0.980	45957.524	0.695	4015.781	74.841	4.457	355.283
359	158.521	1.041	47011.164	0.799	3838.598	70.321	6.975	221.982
360	118.308	1.098	45877.686	0.619	2904.193	45.196	4.997	287.761
361	152.132	1.148	46181.700	0.741	3642.264	73.073	5.230	309.247
362	154.648	0.609	45448.436	0.975	4623.963	58.933	5.424	202.033
363	141.989	0.307	46417.324	1.011	3540.290	54.896	4.715	276.857
364	208.848	0.736	44211.510	0.726	4186.618	66.220	5.849	347.204
365	159.855	0.353	46876.579	0.927	4144.932	65.600	5.129	258.405
366	135.328	0.770	47541.378	0.594	2891.173	41.005	4.707	271.296
367	202.667	0.779	42160.190	0.821	3221.300	68.322	6.130	342.654
368	198.940	0.388	47814.976	0.860	4105.352	90.321	4.895	351.170
369	128.487	0.597	45106.288	0.558	3834.936	34.146	5.259	203.864
370	141.407	0.466	49330.722	0.803	3369.432	71.177	4.603	216.821
371	134.797	0.733	48735.139	0.688	3991.723	51.070	4.800	201.922
372	157.574	0.983	48969.560	0.736	3073.670	68.483	6.080	314.213
373	149.457	2.043	49348.024	0.731	3673.191	72.494	5.007	219.031
374	151.085	0.228	46579.613	0.761	3433.845	55.540	4.413	261.585
375	136.084	0.376	49560.473	0.829	3597.226	56.269	3.892	287.591
376	138.584	0.539	44051.306	0.820	4069.137	66.130	5.270	365.076
377	180.184	0.710	41871.968	0.785	3137.747	74.436	3.711	268.443
378	135.045	0.797	45184.789	0.698	3807.046	46.302	5.594	265.148
379	197.853	0.876	36423.955	0.558	4408.445	72.165	4.453	324.199
380	170.497	0.494	46684.940	0.830	3050.593	60.272	4.924	359.582
381	184.213	0.676	46836.820	0.913	4347.335	97.822	4.790	330.986
383	150.445	0.623	39560.958	0.685	2399.108	43.035	5.266	345.747
384	227.523	0.883	45081.510	0.760	4528.491	99.760	4.273	194.849
385	107.199	0.558	45437.196	0.854	3033.971	49.693	4.792	206.057
386	184.141	0.112	46506.379	0.608	4164.935	52.018	4.097	199.909
387	167.133	0.379	45328.633	0.863	3247.049	34.857	3.789	317.652
388	157.255	0.584	42934.795	0.636	3463.881	53.047	5.251	242.007
389	179.057	0.519	46683.971	0.839	3757.165	105.886	3.831	290.058
390	186.419	0.450	43010.031	0.963	4276.789	83.787	3.937	318.445
391	193.664	0.660	46145.221	0.816	3931.912	77.766	3.871	267.207
392	166.156	0.907	49334.827	0.657	3881.875	60.409	4.009	245.992
393	133.142	0.939	42111.755	0.636	1645.877	38.202	3.762	261.892

Supplementary Table 3.6 (continued)

NSFTV_ID	Fe	I	K	Li	Mg	Mn	Mo	Na
394	230.589	0.660	45401.532	1.057	4358.544	71.465	4.742	354.503
395	155.230	0.490	43450.534	0.834	4351.150	62.876	4.167	250.222
396	165.153	0.047	46081.448	0.742	4000.438	57.256	5.409	265.282
397	141.977	0.145	46562.320	0.984	4014.542	66.184	5.010	299.617
398	125.325	0.710	44595.695	0.765	3671.661	47.514	2.549	341.907
399	210.243	0.787	45716.143	0.822	4428.526	88.354	4.911	308.252
400	112.148	0.753	46047.109	0.708	3411.515	40.731	2.851	220.532

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
1	0.284	7001.397	0.567	5.312	4566.797	18.367	778.300	0.801	292.406	1.479
2	0.187	5875.543	0.463	4.054	4884.523	20.197	703.078	0.801	306.164	1.343
3	0.227	6278.006	0.486	0.296	3906.517	22.047	573.699	0.785	175.274	1.950
4	0.319	5647.216	0.483	4.496	4018.060	23.519	781.878	1.167	361.818	1.105
5	0.409	6029.648	0.498	1.825	4301.577	24.188	748.049	0.781	258.065	1.521
6	0.295	6247.445	0.413	6.015	3913.763	24.041	921.051	1.043	247.906	1.260
7	0.230	6393.031	0.467	1.588	4547.541	26.236	828.434	0.790	418.027	0.862
8	0.218	6065.297	0.428	4.486	4643.266	21.023	674.916	0.631	321.429	1.696
9	0.233	8650.009	0.337	3.160	4484.660	24.187	624.053	0.660	330.423	1.614
10	0.220	6624.886	0.417	2.793	4162.639	23.084	656.168	0.727	223.832	0.895
11	0.218	5446.492	0.538	6.720	4602.919	17.114	779.879	0.783	322.840	1.581
12	0.335	5789.498	0.425	4.671	4331.856	21.928	623.563	0.745	207.026	1.846
13	0.276	7256.376	0.473	6.096	4275.927	21.911	637.504	0.685	281.778	2.905
14	0.232	6630.603	0.415	6.370	5335.402	22.004	731.585	0.852	264.839	1.535
15	0.232	5963.546	0.459	1.100	4219.120	21.798	672.430	0.678	378.720	0.878
16	0.325	6426.801	0.596	2.286	6421.327	34.547	797.817	0.870	299.905	0.955
17	0.216	5131.859	0.513	4.861	4527.938	25.556	646.846	0.771	241.889	1.123
18	0.320	5436.881	0.519	6.301	6009.182	28.333	707.316	0.838	411.262	2.288
19	0.343	5848.329	0.410	3.152	3921.693	22.439	581.572	0.958	433.220	1.928
20	0.316	7241.988	0.485	5.472	4380.829	24.401	669.787	0.709	299.245	1.207
21	0.233	5143.629	0.471	6.123	3568.596	26.890	686.167	0.820	228.592	1.360
22	0.127	7010.533	0.534	4.271	4747.023	22.672	670.492	0.829	339.465	1.779
23	0.192	5764.569	0.474	4.945	4995.863	21.921	549.212	0.937	257.049	2.067
24	0.232	6874.654	0.554	3.476	4784.044	27.748	602.025	0.768	282.915	1.564
25	0.213	6400.097	0.493	1.198	4602.700	27.477	560.956	0.941	299.446	1.439
26	0.198	7512.363	0.598	5.724	4986.740	29.054	525.480	0.715	272.253	1.458
27	0.068	6716.337	0.542	5.474	5424.984	23.944	803.839	0.671	227.272	1.618
28	0.257	5876.974	0.528	1.677	4655.209	23.496	670.979	0.871	274.154	2.872
29	0.201	5617.467	0.671	4.584	4500.569	34.673	431.232	0.966	392.142	1.243
30	0.221	5233.507	0.585	6.058	4704.838	25.265	801.925	0.702	339.509	1.895
31	0.255	6865.948	0.490	3.807	3999.628	23.853	635.936	0.660	319.176	1.039
32	0.233	7498.288	0.398	2.894	3779.288	21.092	763.040	0.723	327.405	0.904
33	0.250	6530.307	0.566	6.193	4546.395	25.016	732.472	0.959	368.380	1.644
34	0.188	5244.858	0.548	6.351	4449.000	24.180	683.049	0.646	306.977	2.204
35	0.186	5057.039	0.580	4.217	4487.972	30.655	702.093	0.840	337.093	2.262
36	0.239	6078.898	0.523	1.937	5119.600	24.177	613.685	0.734	255.745	1.160
37	0.244	6264.055	0.484	7.972	4765.048	21.283	647.174	0.722	302.088	0.897
38	0.268	6618.065	0.535	4.518	4495.997	33.274	663.859	0.772	348.432	1.206
39	0.268	6420.370	0.500	6.732	4748.678	29.862	547.504	0.947	351.637	3.263
40	0.242	7374.033	0.523	3.693	6251.103	22.375	722.194	0.869	314.502	1.718
41	0.275	7432.567	0.486	3.643	5320.786	22.435	686.031	0.695	366.286	1.301
43	0.184	5285.277	0.530	6.892	4145.580	28.916	591.096	0.963	288.375	1.238
44	0.297	6135.726	0.557	2.876	4508.174	26.213	438.210	0.930	313.504	1.774

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
45	0.292	5311.900	0.551	2.440	5046.727	24.081	660.433	0.826	224.918	2.034
46	0.285	7079.893	0.458	2.008	4492.165	24.345	645.921	0.702	304.171	2.657
48	0.201	6037.157	0.576	4.720	5219.949	23.373	623.444	0.754	341.057	1.096
49	0.345	5686.997	0.597	3.369	4384.756	26.828	730.462	0.964	322.196	1.165
50	0.340	6637.400	0.290	6.568	5073.154	24.499	805.302	0.956	276.673	2.118
51	0.253	6959.481	0.529	0.585	4126.478	23.229	746.028	0.705	431.500	0.939
52	0.235	6560.987	0.479	1.599	4128.500	19.629	774.759	0.611	337.749	1.210
53	0.278	6037.752	0.533	3.694	6847.625	27.397	693.183	0.826	303.123	1.618
54	0.218	7860.076	0.525	3.826	4886.069	31.469	1057.317	1.017	280.007	1.022
55	0.235	7407.894	0.492	1.163	4936.004	25.675	624.121	0.844	340.281	1.581
56	0.217	6224.699	0.421	3.319	3914.016	20.691	687.654	0.566	306.725	1.123
57	0.193	5523.310	0.546	3.582	5106.190	28.485	816.053	0.803	430.678	1.941
58	0.241	6639.036	0.572	2.886	4754.575	23.624	597.590	0.792	274.042	1.538
59	0.220	6683.748	0.472	4.059	5120.869	24.441	668.620	0.771	247.909	2.164
60	0.310	6334.057	0.479	3.706	4311.411	21.241	694.304	0.723	224.370	1.030
61	0.184	5475.356	0.533	5.729	4061.688	23.457	685.922	0.791	274.577	2.069
62	0.217	6236.320	0.425	4.382	4447.037	21.645	743.147	0.583	291.499	1.237
63	0.239	6583.658	0.510	4.667	4187.017	21.538	769.473	0.679	386.412	1.248
64	0.229	7138.567	0.478	1.587	4874.420	19.481	705.007	0.817	265.912	0.858
65	0.230	7715.393	0.559	3.894	5311.394	27.366	539.397	0.735	278.280	1.512
66	0.167	5949.451	0.446	5.833	5082.892	24.690	570.267	0.853	255.867	1.353
67	0.219	6596.518	0.511	2.981	4661.035	25.457	683.955	0.860	345.395	1.283
68	0.292	6782.759	0.546	4.672	5991.683	32.413	836.394	1.192	290.841	1.040
69	0.167	5415.911	0.540	2.675	4360.893	23.037	684.204	0.777	367.299	2.229
70	0.220	6679.403	0.505	4.110	5320.713	27.294	744.482	0.721	335.739	1.478
71	0.232	5911.454	0.461	2.752	4667.479	25.775	507.204	0.752	250.383	1.687
72	0.247	5264.121	0.555	4.801	5326.158	28.628	768.102	0.884	320.407	1.180
73	0.191	6261.202	0.464	5.369	4393.582	29.274	609.396	0.716	301.562	1.706
74	0.203	5651.717	0.506	5.768	5193.869	18.995	602.207	0.819	281.258	1.834
75	0.209	7735.375	0.499	3.966	5258.137	25.521	565.067	0.894	352.680	2.425
76	0.211	5857.008	0.594	3.804	4557.648	27.188	765.235	0.848	327.285	1.558
77	0.229	7327.218	0.592	6.420	5629.839	24.161	605.648	0.818	239.019	2.137
78	0.207	5070.945	0.563	4.042	3910.409	25.838	659.005	0.911	340.799	1.789
79	0.246	6392.943	0.539	0.535	4543.958	23.414	812.841	0.689	343.917	1.169
80	0.267	7195.397	0.529	4.056	5542.102	26.659	848.162	1.102	312.341	0.741
81	0.238	5003.586	0.501	3.634	4022.982	21.294	673.161	0.818	357.209	2.036
83	0.249	5930.504	0.545	2.970	4275.491	20.678	727.134	0.620	289.488	0.729
84	0.192	6703.184	0.567	4.805	4743.735	18.433	703.697	0.628	291.195	1.559
85	0.314	5713.792	0.522	2.485	4045.639	22.392	996.505	0.694	212.448	1.258
86	0.276	6202.466	0.520	5.221	3525.578	19.616	568.181	0.591	254.768	2.168
87	0.219	6465.265	0.604	3.953	5049.270	29.957	549.763	0.864	351.351	1.020
88	0.244	6361.404	0.477	5.302	4474.913	23.025	579.341	0.975	334.562	1.374
89	0.221	7189.954	0.503	5.417	4779.955	30.639	598.901	0.992	366.792	1.315
90	0.162	4215.975	0.629	2.999	3771.382	22.546	484.115	0.967	196.611	2.421

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
91	0.167	6476.503	0.525	0.905	4426.556	19.532	696.904	0.689	424.796	0.927
92	0.236	6609.755	0.410	4.038	4065.655	19.603	824.507	0.855	230.029	1.837
93	0.310	5823.640	0.508	3.878	4600.662	22.402	866.461	0.660	216.011	1.764
94	0.258	7912.398	0.502	1.657	3937.903	21.988	813.942	0.773	355.200	0.885
95	0.212	6044.419	0.443	4.339	5094.603	21.883	729.026	0.837	258.467	3.210
96	0.269	6090.639	0.605	3.853	4944.387	18.655	678.929	0.863	306.504	0.991
97	0.237	4778.910	0.577	2.954	3716.570	24.167	733.274	0.916	372.976	1.373
98	0.175	6323.144	0.630	-0.554	4826.059	21.626	709.109	0.746	244.875	0.584
99	0.151	6053.532	0.510	1.641	3579.206	20.908	620.108	0.675	323.094	1.874
100	0.199	6424.131	0.492	4.110	4093.808	23.243	755.355	0.720	283.497	1.501
101	0.142	9285.017	0.578	5.442	4607.933	18.543	632.472	0.681	231.652	1.319
102	0.166	5218.042	0.558	4.863	4358.432	29.817	791.218	1.045	248.794	1.447
103	0.167	7296.679	0.610	2.628	5049.545	27.823	743.701	0.806	249.604	0.783
104	0.183	7049.793	0.612	1.864	3999.292	19.496	754.181	0.858	247.352	0.951
105	0.179	6502.527	0.616	3.272	4386.606	25.050	790.001	1.101	336.216	0.939
106	0.089	4720.098	0.590	5.680	3694.386	22.903	551.171	0.654	169.517	2.442
107	0.088	6041.879	0.537	5.844	4644.724	22.358	718.762	0.805	209.612	1.457
108	0.182	7519.500	0.504	4.812	5016.180	16.882	672.242	0.659	202.473	1.519
109	0.208	6180.346	0.619	6.615	3571.584	25.073	549.509	0.710	286.132	1.527
110	0.138	6258.557	0.611	3.344	4785.345	19.156	524.544	0.946	253.995	2.814
111	0.127	6011.106	0.593	4.070	4816.019	23.449	766.812	0.772	314.210	1.622
112	0.274	6239.167	0.684	2.498	4258.290	19.792	610.437	0.762	152.183	1.592
113	0.193	6911.108	0.604	2.048	3737.053	19.425	636.164	0.539	378.355	1.203
114	0.126	5600.366	0.582	2.415	3539.662	23.197	832.553	0.625	198.991	1.498
115	0.106	5863.033	0.555	1.125	4244.020	15.452	632.417	0.641	228.771	1.686
116	0.191	6421.851	0.581	3.234	5053.146	27.501	742.909	0.750	328.553	1.185
117	0.118	4328.630	0.692	2.958	3798.866	23.388	697.318	1.073	228.490	1.565
118	0.193	6008.636	0.566	3.837	4074.624	18.020	715.760	0.731	223.362	1.401
119	0.148	5185.455	0.486	4.145	4486.533	20.263	623.288	0.759	285.541	1.849
120	0.112	5709.585	0.558	2.158	4287.375	26.029	736.617	0.979	391.950	1.480
121	0.149	6825.980	0.570	0.475	4255.755	23.309	769.108	0.735	290.859	1.008
122	0.124	6366.110	0.721	3.247	5313.022	23.713	678.425	0.988	250.820	2.477
123	0.148	4629.696	0.593	3.042	4231.001	26.497	699.504	0.870	405.801	1.383
124	0.213	5194.907	0.678	-1.394	7017.344	22.248	768.645	0.940	226.578	1.915
125	0.124	5662.400	0.661	3.318	4541.213	23.016	630.477	0.788	200.877	2.161
126	0.073	5673.577	0.565	3.691	4779.288	22.063	701.244	0.556	332.630	0.922
127	0.098	7610.411	0.568	1.170	5219.959	28.271	566.525	0.836	406.590	0.721
128	0.140	7512.311	0.717	0.322	4840.384	25.716	751.925	0.651	280.061	0.839
129	0.121	5733.003	0.600	2.704	5172.547	27.247	621.996	0.889	292.334	1.443
130	0.248	4432.364	0.645	3.768	4234.793	27.986	679.089	0.823	323.937	1.485
131	0.243	5134.575	0.628	3.336	4187.108	26.103	507.723	1.050	371.050	1.469
132	0.132	5614.600	0.631	3.229	4772.246	19.175	640.941	0.801	276.604	1.623
133	0.160	7126.176	0.651	-1.489	4034.501	20.864	669.500	0.687	308.418	0.751

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
134	0.224	6173.030	0.616	3.027	4366.871	21.951	670.198	0.690	318.903	1.916
135	0.080	6132.755	0.645	3.157	4555.884	23.290	769.298	0.830	319.198	1.438
136	0.200	4678.424	0.561	4.738	3714.285	26.382	551.176	0.819	191.973	2.134
137	0.148	5637.599	0.604	3.206	5419.483	23.872	554.666	0.712	281.820	2.452
138	0.184	4686.418	0.519	2.523	5110.514	26.159	525.695	0.742	235.458	2.125
139	0.221	6863.535	0.570	4.265	4343.020	25.445	670.321	0.898	333.122	1.053
140	0.177	6184.058	0.680	3.977	4415.928	25.388	765.825	0.781	349.459	0.781
141	0.147	6442.850	0.536	2.882	5212.663	25.802	560.180	0.921	312.720	0.977
142	0.194	6850.392	0.468	3.496	4901.948	29.145	747.719	0.784	329.210	1.196
143	0.145	7004.095	0.577	2.274	3980.663	22.028	808.563	0.695	315.744	0.942
144	0.113	7365.843	0.509	1.679	4305.619	24.708	844.577	0.771	241.771	0.745
145	0.146	7092.090	0.593	3.832	5857.563	20.199	721.385	0.747	311.239	1.470
146	0.195	5460.826	0.554	3.810	4709.790	25.960	644.502	0.850	307.512	1.941
147	0.187	7502.089	0.660	0.608	4656.718	23.683	801.064	0.871	283.917	1.210
148	0.133	5851.734	0.673	1.020	4465.407	23.622	600.570	0.745	285.951	2.424
149	0.125	6261.757	0.564	2.227	3725.910	21.303	724.818	0.845	310.849	1.346
150	0.112	6841.349	0.515	2.034	4365.057	20.433	640.964	0.705	262.426	1.428
151	0.201	6331.718	0.661	1.804	4674.111	24.266	650.080	0.650	299.808	0.852
152	0.236	5361.316	0.492	5.769	4465.873	26.462	773.067	1.012	291.733	1.358
153	0.171	5006.076	0.615	2.598	4350.013	25.883	615.230	1.166	350.228	2.071
154	0.181	9580.677	0.495	3.152	7108.141	20.221	610.019	0.618	303.266	2.732
155	0.191	9284.843	0.613	1.548	6722.395	22.154	595.446	0.701	299.133	1.968
156	0.180	5305.283	0.535	2.407	4487.424	31.378	692.533	0.918	361.164	1.389
157	0.138	5517.264	0.542	-0.882	4223.624	21.975	544.675	0.690	284.006	0.833
158	0.092	6894.163	0.659	-2.971	4094.383	21.705	642.060	0.678	319.233	0.971
159	0.169	4562.428	0.656	4.175	3414.624	21.958	636.015	0.970	257.944	2.401
160	0.197	4695.685	0.642	0.781	6177.065	24.586	656.979	0.703	258.832	1.442
161	0.185	5350.999	0.719	2.518	4303.950	31.637	775.818	0.893	464.420	0.918
162	0.141	4732.141	0.663	3.944	5274.457	28.639	648.570	0.911	251.568	1.728
163	0.167	5429.424	0.678	3.594	5916.556	26.213	689.561	0.684	269.419	2.058
164	0.236	7501.523	0.515	4.155	5399.707	22.306	853.843	0.719	316.309	1.273
165	0.086	7822.851	0.585	4.884	5704.449	21.653	735.021	0.714	314.159	1.634
166	0.207	6701.220	0.734	5.856	4141.849	26.420	629.784	0.940	272.615	1.341
167	0.153	7536.506	0.578	3.488	5229.146	25.112	498.075	0.740	298.468	1.160
168	0.226	7281.382	0.539	2.396	4278.821	25.897	856.402	1.335	209.474	0.915
169	0.123	7819.075	0.683	1.615	4029.487	22.877	765.639	0.594	313.629	1.454
170	0.218	7030.066	0.575	4.943	4746.090	26.144	680.757	0.790	235.075	1.675
171	0.119	5292.898	0.560	2.333	4628.331	25.268	565.431	0.904	309.599	2.089
172	0.122	5396.383	0.542	5.473	4266.175	24.203	720.074	0.818	204.960	1.700
173	0.178	6908.962	0.604	1.686	4490.620	22.654	601.701	0.768	283.515	0.612
174	0.212	7528.874	0.438	4.657	5307.479	27.198	588.168	0.889	270.588	1.514
175	0.227	7347.480	0.564	5.127	5492.849	26.282	595.608	0.829	262.285	1.754
176	0.257	6436.625	0.419	4.459	4067.853	27.982	812.918	1.001	240.879	0.613

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
177	0.247	6635.035	0.465	4.185	5006.426	23.727	717.433	0.844	212.290	0.718
178	0.349	6464.101	0.439	4.172	4893.977	23.563	871.727	1.043	306.867	0.616
179	0.313	7426.877	0.377	5.622	4448.477	26.347	670.469	0.703	317.643	1.164
180	0.229	6181.452	0.246	3.957	3514.480	21.095	675.618	0.789	256.436	1.170
181	0.265	6921.143	0.284	3.864	3772.108	23.550	597.121	0.868	276.531	1.346
182	0.294	6882.843	0.329	5.338	4998.549	27.686	715.321	0.824	300.564	1.412
183	0.235	7385.560	0.366	3.809	4703.415	28.849	622.981	0.896	332.651	1.082
184	0.284	7451.847	0.223	2.235	4729.612	23.822	602.846	0.979	346.244	1.865
185	0.266	7041.826	0.338	1.979	4509.091	29.837	730.990	0.967	263.246	1.579
186	0.207	6298.631	0.443	2.726	4384.376	25.579	936.133	0.906	293.007	0.360
187	0.157	7835.541	0.375	3.319	4922.211	28.682	787.283	0.776	276.872	0.823
188	0.195	8148.998	0.267	3.215	4623.129	25.559	857.956	0.924	259.000	0.703
189	0.207	5635.499	0.322	4.282	4866.616	26.910	632.465	1.103	374.037	1.984
190	0.157	7063.364	0.375	1.712	4387.269	26.179	571.946	0.918	252.263	1.121
191	0.294	7235.165	0.318	2.128	6172.775	29.363	716.353	0.963	319.504	1.934
192	0.270	9279.608	0.260	4.818	5768.054	24.660	670.092	0.862	303.099	1.181
193	0.231	7067.486	0.441	3.948	4716.383	25.436	751.085	0.847	242.912	1.279
194	0.169	6337.513	0.255	0.171	5186.823	20.513	576.934	1.037	260.414	3.886
195	0.122	7428.044	0.284	4.627	5375.405	28.483	665.074	0.837	238.936	1.358
196	0.159	6641.860	0.283	5.894	4809.281	22.495	608.434	0.889	287.720	2.215
197	0.266	9298.326	0.355	1.838	5939.756	25.372	578.181	1.009	302.829	1.472
198	0.139	6265.966	0.459	1.015	4444.710	28.062	562.258	0.997	270.098	0.806
199	0.178	6989.836	0.247	3.181	5106.007	27.204	608.119	0.883	202.490	1.392
200	0.257	7516.820	0.266	3.719	5888.061	24.117	587.326	1.028	340.762	2.756
201	0.125	7630.167	0.338	3.465	4963.336	25.744	724.109	0.795	329.534	1.300
202	0.212	6650.821	0.304	1.727	5004.304	32.497	638.536	0.945	298.378	1.590
203	0.177	7577.372	0.279	5.064	5066.681	27.371	602.050	0.864	295.518	2.103
204	0.209	7150.404	0.454	4.103	4815.264	21.075	496.681	0.717	358.394	1.351
205	0.271	7549.938	0.276	4.613	6192.602	26.756	630.033	1.158	318.863	1.483
206	0.266	8069.900	0.286	3.593	4230.224	27.575	663.823	1.416	231.559	1.278
207	0.138	5836.292	0.365	2.501	4420.426	24.637	548.049	0.931	194.337	1.616
208	0.163	6839.536	0.404	5.392	5863.197	28.318	554.145	0.944	326.790	2.374
209	0.187	5974.439	0.296	6.685	4494.528	28.212	506.839	0.837	226.909	1.533
211	0.245	7365.835	0.372	3.559	5508.730	30.239	760.845	0.953	317.717	0.893
212	0.189	7575.229	0.362	3.895	4498.558	27.603	637.429	0.947	282.932	1.352
213	0.159	8638.706	0.378	2.959	5634.312	31.226	681.923	0.951	307.534	1.790
214	0.210	7329.394	0.220	5.422	4778.392	28.049	581.278	0.876	283.399	1.256
215	0.153	8044.355	0.335	1.870	5220.953	29.222	732.446	1.009	287.813	1.201
216	0.238	7233.976	0.373	3.782	4897.381	30.174	682.655	0.977	292.600	0.937
217	0.290	5944.792	0.394	4.408	4574.083	22.159	740.866	0.872	229.435	1.204
218	0.212	9167.205	0.355	2.984	5149.709	29.637	634.372	0.962	307.491	0.837
219	0.266	8789.339	0.313	2.647	5701.787	22.825	618.618	0.702	314.587	2.381
220	0.163	9012.277	0.237	3.278	5113.465	22.406	534.718	0.669	323.276	2.384

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
221	0.327	5380.332	0.480	2.639	5562.993	26.996	664.179	0.856	255.280	2.334
222	0.211	6701.523	0.318	6.676	5606.647	22.417	567.180	0.774	298.482	2.324
223	0.130	6608.812	0.434	2.965	4620.109	32.127	711.198	1.013	269.273	1.832
224	0.184	6876.976	0.288	1.117	4628.721	20.689	653.815	0.935	231.921	2.569
225	0.181	7456.417	0.309	1.899	4183.563	18.603	584.030	0.699	196.094	1.574
226	0.151	6459.145	0.413	3.153	4365.572	20.633	628.792	0.843	209.114	2.069
227	0.225	7469.823	0.335	2.731	5212.876	25.916	714.481	0.918	266.269	1.442
228	0.268	6295.376	0.189	3.886	5823.630	22.221	544.162	1.002	335.974	2.377
229	0.238	6640.773	0.277	1.494	4236.399	20.257	641.469	0.773	254.520	1.410
230	0.182	7392.031	0.327	2.548	4726.739	20.867	582.952	0.829	223.614	1.565
231	0.224	6266.194	0.396	3.674	4833.503	28.486	602.957	0.962	311.314	2.000
232	0.203	6681.366	0.329	3.144	4552.476	30.345	688.603	0.987	315.849	0.578
233	0.209	6599.690	0.340	1.610	4146.768	25.242	606.101	0.724	386.743	1.475
234	0.157	6506.615	0.388	2.625	5216.556	27.189	545.623	0.942	290.850	2.101
235	0.165	5822.618	0.415	2.233	5466.326	28.964	548.931	0.940	363.522	2.171
236	0.218	6054.905	0.371	2.174	3668.325	23.535	546.792	0.842	333.756	0.974
237	0.230	6227.633	0.404	-0.633	4522.838	20.083	679.295	0.921	268.714	0.762
238	0.133	6945.227	0.380	3.794	4835.152	26.519	668.969	0.978	220.282	1.043
239	0.281	6494.437	0.363	1.642	4628.647	25.953	636.153	0.839	232.127	1.418
240	0.264	6109.225	0.396	4.261	4775.521	26.890	630.573	0.941	283.156	1.585
241	0.225	6774.983	0.230	1.810	4806.984	25.594	594.101	1.015	288.181	2.183
242	0.175	7750.916	0.431	2.465	5226.874	32.382	823.103	0.940	253.264	1.346
243	0.162	8082.079	0.318	2.109	5008.714	20.732	638.292	0.816	363.308	2.326
244	0.246	6625.751	0.239	4.792	4938.524	23.536	739.106	0.795	213.888	1.311
245	0.191	7541.074	0.247	2.672	4639.338	26.521	541.853	0.901	317.115	2.381
246	0.198	7612.639	0.322	3.641	4684.455	26.022	535.990	1.023	321.509	2.443
247	0.198	6806.872	0.365	3.045	5393.519	26.385	700.479	0.854	311.543	1.181
248	0.175	7028.393	0.309	2.139	4455.620	26.485	622.137	0.866	346.740	0.935
249	0.167	6624.904	0.263	3.018	4802.064	24.450	728.178	0.940	187.457	1.512
250	0.172	7412.776	0.400	2.970	3882.473	25.271	640.990	0.693	260.424	1.234
251	0.205	6767.228	0.417	6.167	3739.427	25.693	673.766	0.916	192.578	0.965
252	0.171	5656.517	0.238	3.174	3859.481	23.632	641.845	0.743	234.519	2.616
253	0.252	7543.018	0.367	2.666	5036.317	22.559	785.796	0.919	230.581	0.791
254	0.214	8313.210	0.229	3.643	3985.969	23.072	660.504	0.968	281.817	1.324
255	0.153	6596.617	0.519	5.058	5585.952	28.146	577.783	0.907	334.197	3.217
256	0.162	7005.606	0.174	1.129	4843.756	19.470	627.809	0.805	281.427	2.340
257	0.189	6434.275	0.335	1.267	5115.678	18.556	616.577	0.823	247.469	2.180
258	0.186	6795.108	0.338	4.302	4750.672	25.353	775.956	0.874	319.775	2.316
259	0.227	6609.952	0.206	3.676	4834.258	23.235	700.933	1.144	265.702	2.028
260	0.312	5817.221	0.354	3.258	5683.168	29.341	707.752	0.977	262.789	1.799
261	0.321	5879.029	0.305	3.300	4154.797	24.222	653.407	1.241	237.658	2.028
262	0.337	6148.421	0.450	3.966	4650.503	23.388	537.420	0.927	301.316	3.056
263	0.291	7363.963	0.344	4.308	4225.940	22.372	682.374	0.872	271.825	1.135
264	0.283	7420.766	0.184	2.857	4908.344	25.602	475.614	0.866	250.216	1.630

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
265	0.279	7648.411	0.203	1.810	5343.182	23.875	601.334	0.841	347.224	1.552
266	0.264	7066.431	0.284	3.039	4346.507	24.227	735.098	0.782	419.410	1.158
267	0.222	7161.912	0.276	1.803	3708.477	20.578	652.081	0.726	354.312	0.921
268	0.219	7600.917	0.320	4.179	5105.151	20.323	565.577	0.758	326.074	2.392
269	0.202	5628.514	0.265	4.010	4148.566	28.801	607.925	0.824	200.099	1.608
270	0.245	7730.546	0.311	1.125	5406.013	24.155	566.896	0.900	346.541	1.376
271	0.268	6696.597	0.283	2.027	4621.470	24.206	454.795	0.959	240.429	0.921
272	0.249	6607.991	0.342	1.686	4510.348	34.221	669.059	0.991	424.344	1.359
273	0.205	7516.801	0.197	1.825	5925.374	20.003	571.827	0.890	223.277	1.942
274	0.199	8249.198	0.396	2.898	4348.650	24.795	688.147	0.834	286.165	1.690
275	0.246	7165.066	0.257	2.492	4874.574	22.813	650.083	0.692	346.919	1.381
276	0.288	6991.564	0.290	4.865	6163.385	31.799	750.963	1.253	296.136	1.213
277	0.217	7407.679	0.406	3.173	4511.308	25.856	681.337	0.819	297.025	0.751
278	0.288	6794.500	0.291	5.313	3616.356	21.714	695.046	0.809	355.237	0.778
279	0.221	7568.782	0.341	2.531	4888.311	27.307	486.733	0.848	324.556	2.053
280	0.215	7039.097	0.308	2.335	4722.174	21.651	605.085	0.870	239.295	1.211
281	0.218	6882.252	0.207	2.337	4694.899	26.191	697.661	0.800	227.031	1.313
282	0.236	7171.477	0.314	3.101	4030.708	23.352	795.086	0.789	355.173	0.849
283	0.225	7498.213	0.265	4.529	4004.542	25.629	695.841	0.824	401.031	0.879
284	0.256	6056.082	0.381	3.551	4992.706	27.753	639.092	0.819	270.335	2.430
285	0.223	6278.334	0.223	5.045	4668.138	25.750	767.667	1.003	262.299	0.899
286	0.198	6995.955	0.377	2.143	4422.083	25.392	704.440	0.936	247.950	1.403
287	0.196	7728.453	0.294	2.971	3976.736	19.579	594.702	0.750	301.182	2.517
288	0.270	7200.641	0.340	2.967	4632.902	23.540	653.560	0.844	409.783	1.180
289	0.201	7313.069	0.435	2.986	4824.851	24.072	619.032	0.849	452.268	1.504
290	0.261	6457.735	0.472	4.004	4011.740	20.893	756.619	0.793	281.332	0.848
291	0.173	7502.576	0.319	4.398	4467.960	23.774	685.804	0.828	255.322	0.858
292	0.212	6945.165	0.338	2.733	4562.005	21.257	593.060	0.735	285.719	2.171
293	0.174	6326.356	0.303	3.069	4574.235	23.765	699.339	0.938	277.224	1.889
294	0.217	6497.349	0.364	4.753	4465.588	26.828	902.397	1.075	219.171	0.644
295	0.232	8176.608	0.372	4.191	4479.942	28.634	757.663	0.842	408.276	1.143
296	0.238	5844.844	0.238	3.008	3993.746	22.785	855.980	0.860	245.918	1.561
297	0.245	7434.873	0.269	3.259	4403.498	24.644	736.984	0.788	314.235	1.168
298	0.198	6564.581	0.307	5.273	4332.495	29.369	666.519	1.004	225.179	1.656
299	0.192	6089.295	0.437	5.127	4552.537	27.939	659.243	0.914	291.558	1.397
300	0.238	6791.287	0.366	4.013	4170.755	25.425	720.343	0.736	373.242	0.903
301	0.211	8171.729	0.262	1.483	5111.326	25.415	634.753	0.934	362.421	2.619
302	0.192	8092.903	0.251	3.531	5166.507	22.106	677.632	0.835	251.149	1.322
303	0.224	7172.144	0.323	3.708	5101.492	23.677	710.381	0.892	192.369	1.129
304	0.233	7997.767	0.344	1.791	4723.909	25.745	607.865	0.852	262.295	1.422
305	0.215	6639.161	0.234	3.226	5128.198	23.708	630.643	1.004	292.163	1.220
306	0.210	7689.656	0.399	3.035	4674.245	18.238	602.702	0.685	199.209	2.068
307	0.235	8899.651	0.315	2.405	5315.287	22.934	676.697	0.680	337.988	0.918
308	0.186	6546.269	0.309	1.382	5093.115	30.075	652.221	0.957	302.558	1.232

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
309	0.167	6987.441	0.372	2.889	5131.624	28.235	591.021	0.852	241.915	1.373
310	0.149	6782.777	0.391	3.294	4765.148	31.710	776.263	0.913	255.941	0.847
311	0.246	7279.778	0.345	3.327	4514.915	24.997	755.152	0.667	409.846	0.416
312	0.295	5244.843	0.309	3.082	3635.764	27.700	601.683	1.026	286.203	1.844
313	0.156	6442.315	0.382	5.187	5291.502	27.747	531.930	0.981	363.356	1.806
314	0.276	6117.724	0.344	4.667	4691.912	22.513	744.694	0.972	279.547	1.550
315	0.174	6891.844	0.224	-1.247	5235.556	23.561	565.440	0.932	282.131	2.941
316	0.271	6523.799	0.295	3.713	4817.011	25.964	674.385	1.184	346.641	1.686
317	0.239	6801.580	0.223	1.813	4244.510	23.965	771.794	1.116	274.569	1.940
318	0.311	5325.839	0.288	2.226	3958.032	25.772	667.032	0.974	292.671	1.722
319	0.286	6452.040	0.248	2.951	4716.534	27.674	691.066	1.132	357.186	1.737
320	0.244	6058.246	0.262	0.896	4354.723	25.673	658.524	0.956	335.330	1.247
321	0.288	6884.441	0.252	2.142	5365.164	32.659	696.204	1.130	351.772	1.485
322	0.212	6184.339	0.272	2.209	4874.493	25.580	707.029	1.057	258.720	1.695
323	0.259	7405.035	0.244	1.467	4499.472	25.054	621.925	0.852	307.103	2.287
324	0.284	6325.562	0.351	2.928	5221.899	30.042	600.507	1.060	317.202	1.933
325	0.163	6576.540	0.179	0.801	5916.370	23.358	710.588	0.893	205.404	1.484
326	0.279	6190.549	0.305	1.483	4283.547	31.643	637.452	1.011	328.581	1.624
327	0.299	6391.017	0.303	3.548	4721.027	26.860	614.548	0.957	300.213	2.547
328	0.250	5705.388	0.313	1.927	4995.394	25.904	756.143	0.894	378.702	1.658
329	0.248	7155.008	0.235	3.727	4030.561	21.908	570.844	1.007	267.050	2.168
330	0.276	6793.117	0.416	3.161	5204.836	26.695	681.950	0.962	374.753	1.618
331	0.226	5834.057	0.270	3.870	5476.253	21.386	667.891	0.857	253.181	1.910
332	0.208	4997.074	0.264	1.741	5145.598	30.531	783.297	1.218	187.012	0.922
333	0.250	8464.158	0.326	3.433	4063.903	24.860	535.592	0.705	339.698	0.737
334	0.219	6524.466	0.308	2.328	4185.318	26.807	850.400	0.898	237.122	1.314
335	0.245	7237.355	0.337	3.486	5169.699	30.058	756.584	0.882	313.291	1.080
336	0.303	5676.415	0.301	2.956	3973.206	22.847	631.162	0.996	356.571	1.996
337	0.186	6313.288	0.410	5.004	4405.408	27.447	627.570	0.935	237.611	2.283
338	0.176	7512.967	0.301	1.315	4047.686	25.327	690.705	0.740	312.227	1.127
339	0.199	6370.106	0.371	3.801	5198.164	26.416	608.607	0.863	379.507	2.101
340	0.227	7103.149	0.205	3.704	4841.853	23.359	657.489	0.965	236.820	1.444
341	0.201	6691.748	0.403	4.805	4432.694	23.943	673.041	0.908	260.581	1.945
342	0.178	7636.931	0.367	3.614	4227.216	29.202	803.322	0.874	338.487	1.324
343	0.276	6807.101	0.390	2.957	4682.789	23.999	678.634	0.690	306.297	1.528
344	0.239	5406.481	0.253	5.846	4658.451	21.865	700.459	0.810	229.123	1.156
345	0.317	7696.945	0.383	5.219	5634.206	24.100	755.723	1.117	277.037	2.756
346	0.305	6812.753	0.380	6.462	4960.761	28.653	707.437	0.971	332.422	1.802
347	0.192	8027.308	0.497	4.260	5254.924	27.337	600.043	0.897	268.490	1.779
348	0.352	6800.459	0.398	5.121	4999.050	28.268	627.972	0.913	303.044	2.102
349	0.231	5860.705	0.401	7.450	4683.989	27.341	521.434	1.052	311.270	2.111
350	0.247	6974.497	0.517	5.313	4739.086	25.219	852.567	0.903	328.701	1.334
351	0.320	7115.698	0.412	5.132	4615.142	23.671	674.785	0.903	295.282	0.682
352	0.277	7296.583	0.470	5.450	5177.358	30.386	692.450	0.819	278.603	1.932

Supplementary Table 3.6 (continued)

NSFTV_ID	Ni	P	Pb	Rb	S	Se	Si	Sr	Zn	DW (g)
353	0.338	6647.343	0.426	6.161	4213.434	27.285	580.722	1.045	294.827	1.261
354	0.233	4320.620	0.476	5.644	4342.659	29.270	709.235	0.904	316.555	1.580
355	0.270	6974.911	0.414	4.332	3997.065	20.262	519.769	0.791	336.985	1.417
356	0.134	6989.257	0.385	6.380	4161.672	26.146	561.155	0.861	319.793	1.788
358	0.235	7141.228	0.428	3.647	4342.517	23.232	642.644	0.784	275.416	1.673
359	0.275	6026.063	0.419	4.243	4983.598	24.491	677.333	0.948	363.251	1.340
360	0.293	6615.660	0.445	5.539	5262.649	26.724	370.224	1.083	326.436	2.731
361	0.230	7195.378	0.406	5.734	5184.032	27.163	673.970	0.866	232.444	1.614
362	0.153	7990.347	0.461	4.311	5495.520	32.596	632.560	0.899	228.681	2.297
363	0.271	6737.957	0.504	2.165	3624.048	24.834	497.308	0.720	309.467	1.135
364	0.180	7262.599	0.440	4.583	5521.574	22.456	597.426	0.721	274.309	1.845
365	0.250	7992.971	0.452	3.645	5485.823	25.025	679.781	0.762	207.174	0.456
366	0.196	6345.240	0.406	3.213	3508.565	23.277	481.846	0.669	291.251	0.782
367	0.256	4944.562	0.430	4.763	4394.078	25.429	626.830	0.910	332.874	1.236
368	0.261	5606.233	0.466	3.547	4265.971	26.659	710.577	0.924	307.706	0.678
369	0.205	6621.143	0.451	3.398	4686.893	23.505	453.801	0.977	274.314	5.084
370	0.274	5755.064	0.458	2.960	4554.097	21.533	513.856	0.967	254.900	2.316
371	0.156	6369.456	0.439	3.098	5139.247	32.534	581.218	0.874	340.056	2.592
372	0.198	7136.258	0.465	2.175	4795.275	26.197	596.675	1.007	346.400	1.742
373	0.213	6199.427	0.543	0.949	6799.898	25.449	566.530	0.812	245.951	1.944
374	0.204	4187.841	0.495	2.123	2809.221	26.518	652.970	0.564	312.121	1.269
375	0.150	7453.836	0.452	3.045	4975.200	26.543	556.994	0.761	280.469	2.260
376	0.188	6361.250	0.398	2.477	4108.721	28.575	551.292	0.960	282.613	1.521
377	0.258	5670.172	0.442	4.543	4033.255	30.616	531.620	0.773	303.939	1.320
378	0.275	5147.205	0.418	2.208	4471.184	26.429	636.359	1.297	304.528	2.691
379	0.120	5829.175	0.442	1.289	4563.856	31.063	551.792	0.986	270.560	1.884
380	0.172	7379.730	0.398	1.243	3576.781	26.060	458.161	0.614	435.114	0.731
381	0.200	6513.296	0.446	2.324	4489.502	28.907	749.281	0.800	380.532	1.309
383	0.182	5827.089	0.554	3.503	4104.984	22.593	562.197	0.692	286.402	1.795
384	0.162	7478.743	0.457	4.301	4769.638	35.219	584.544	0.937	312.882	1.517
385	0.154	6114.267	0.613	3.340	4384.952	25.366	555.431	0.840	253.036	2.677
386	0.222	7342.502	0.430	0.054	4940.640	25.326	593.332	0.632	199.361	1.463
387	0.231	7172.723	0.484	2.833	4863.249	25.873	521.623	0.673	333.005	1.113
388	0.197	6823.150	0.461	2.312	4836.304	25.286	569.051	0.715	455.251	0.976
389	0.292	6307.517	0.413	3.825	3920.722	27.744	613.322	0.763	196.026	0.727
390	0.296	4517.816	0.546	4.913	3976.645	28.078	667.989	0.850	301.425	0.834
391	0.164	7399.956	0.479	3.411	4648.772	25.725	695.361	0.695	298.198	1.073
392	0.124	7630.581	0.430	1.560	5000.662	27.994	626.778	0.815	292.767	1.422
393	0.186	6621.694	0.379	4.083	5091.168	29.347	462.830	0.841	301.842	2.187
394	0.142	7424.645	0.535	3.568	5055.148	31.206	659.626	0.928	288.388	1.537
395	0.075	6264.849	0.549	1.123	4116.453	23.983	599.379	0.675	232.240	2.167
396	0.129	7093.786	0.463	3.307	4462.067	24.000	576.383	0.691	306.467	1.423
397	0.143	7411.469	0.455	4.728	4297.996	27.129	592.776	0.651	307.713	1.307
398	0.165	6775.443	0.431	4.034	5104.016	27.912	501.979	0.799	440.932	1.018
399	0.112	6303.097	0.424	5.235	4872.239	25.642	696.593	0.858	294.010	1.746
400	0.099	6507.529	0.452	2.393	4645.737	25.586	548.107	0.725	321.181	2.074

Supplementary Table 3.6 (continued)

REFERENCES

Agrama HA, Yan W, Jia M, Fjellstrom R, McClung AM (2010) Genetic structure associated with diversity and geographic distribution in the USDA rice world collection. Natural Science 2:247-291

Alam S, Rahman MH, Kamei S, Kawai S (2002) Alleviation of manganese toxicity and manganese-induced iron deficiency in barley by additional potassium supply in nutrient solution. Soil Sci Plant Nutr 48:387-392

Antoine W, Stewart JM, De Los Reyes, Benildo G (2005) The rice homolog of the sodium/lithium tolerance gene functions as molecular chaperon in vitro. Physiol Plantarum 125:299-310

Asch F, Dingkuhn M, Dörffling K, Miezan K (2000) Leaf K/Na ratio predicts salinity induced yield loss in irrigated rice. Euphytica 113:109-118

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465:627-631

Barrett J, Fry B, Maller J, Daly M (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265

Baxter I, Hosmani PS, Rus A, Lahner B, Borevitz JO, Muthukumar B, Mickelbart MV, Schreiber L, Franke RB, Salt DE (2009) Root suberin forms an extracellular barrier that affects water relations and mineral nutrition in Arabidopsis. PLoS genetics 5:e1000492

Baxter I, Muthukumar B, Park HC, Buchner P, Lahner B, Danku J, Zhao K, Lee J, Hawkesford MJ, Guerinot ML (2008) Variation in molybdenum content across broadly distributed populations of Arabidopsis thaliana is controlled by a mitochondrial molybdenum transporter (*MOT1*). PLoS Genetics 4:e1000004

Baxter IR, Gustin JL, Settles AM, Hoekenga OA (2013) Ionomic Characterization of Maize Kernels in the Intermated B73× Mo17 Population. Crop Sci 53:208-220

Buescher E, Achberger T, Amusan I, Giannini A, Ochsenfeld C, Rus A, Lahner B, Hoekenga O, Yakubova E, Harper JF (2010) Natural genetic variation in selected populations of Arabidopsis thaliana is associated with ionomic differences. PloS one 5:e11081

Chao D, Silva A, Baxter I, Huang YS, Nordborg M, Danku J, Lahner B, Yakubova E, Salt DE (2012) Genome-wide association studies identify heavy metal ATPase3 as the

primary determinant of natural variation in leaf cadmium in Arabidopsis thaliana. PLoS Genetics 8:e1002923

Chen Z, Watanabe T, Shinano T, Ezawa T, Wasaki J, Kimura K, Osaki M, Zhu Y (2009) Element interconnections in Lotus japonicus: A systematic study of the effects of element additions on different natural variants. Soil Sci Plant Nutr 55:91-101

Chen Z, Watanabe T, Shinano T, Okazaki K, Osaki M (2009) Rapid characterization of plant mutants with an altered ion‐profile: a case study using Lotus japonicus. New Phytol 181:795-801

Cobb JN, DeClerck G, Greenberg A, Clark R, McCouch S (2013) Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. Theor Appl Genet:1-21

Eide DJ, Clark S, Nair TM, Gehl M, Gribskov M, Guerinot ML, Harper JF (2005) Characterization of the yeast ionome: a genome-wide analysis of nutrient mineral and trace element homeostasis in Saccharomyces cerevisiae. Genome Biol 6:R77

Eizenga G, Prasad B, Jackson A, Jia M (2013) Identification of rice sheath blight and blast quantitative trait loci in two different O. sativa/O. nivara advanced backcross populations. Mol Breed:1-19

Famoso AN, Clark RT, Shaff JE, Craft E, McCouch SR, Kochian LV (2010) Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal aluminum tolerance and investigations into rice aluminum tolerance mechanisms. Plant Physiol 153:1678-1691

Famoso AN, Zhao K, Clark RT, Tung C, Wright MH, Bustamante C, Kochian LV, McCouch SR (2011) Genetic architecture of aluminum tolerance in rice (Oryza sativa) determined through genome-wide association analysis and QTL mapping. PLoS genetics 7:e1002221

Fitzgerald MA, McCouch SR, Hall RD (2009) Not just a grain of rice: the quest for quality. Trends Plant Sci 14:133-139

Fuller DQ (2011) Pathways to Asian civilizations: tracing the origins and spread of rice and rice cultures. Rice 4:78-92

Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun G (2009) The domestication process and domestication rate in rice: spikelet bases from the Lower Yangtze. Science 323:1607-1610

Fuller DQ, Harvey E, Qin L (2007) Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. ANTIQUITY-OXFORD- 81:316

Fuller D, Sato Y, Castillo C, Qin L, Weisskopf A, Kingwell-Banham E, Song J, Ahn S, van Etten J (2010) Consilience of genetics and archaeobotany in the entangled history of rice. Archaeological and Anthropological Sciences 2:115-131

Gabriel KR (1971) The biplot graphic display of matrices with application to principal component analysis. Biometrika 58:453-467

Garciadeblas B, Senn ME, Banuelos MA, Rodríguez‐Navarro A (2003) Sodium transport and HKT transporters: the rice model. The Plant Journal 34:788-801

Garris AJ, Tai TH, Coburn J, Kresovich S, McCOUCH S (2005) Genetic structure and diversity in Oryza sativa L. Genetics 169:1631-1638

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York

Ghandilyan A, Ilk N, Hanhart C, Mbengue M, Barboza L, Schat H, Koornneef M, El-Lithy M, Vreugdenhil D, Reymond M (2009) A strong effect of growth medium and organ type on the identification of QTLs for phytate and mineral concentrations in three Arabidopsis thaliana RIL populations. J Exp Bot 60:1409-1425

Goff SA, Ricke D, Lan T, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92-100

Greenberg AJ, Hackett SR, Harshman LG, Clark AG (2011) Environmental and genetic perturbations reveal different networks of metabolic regulation. Mol Syst Biol 7:563

Gu Y, Zhao Z, Pearsall DM (2013) Phytolith morphology research on wild and domesticated rice species in East Asia. Quaternary International 287:141-148

Hänsch R, Mendel RR (2009) Physiological functions of mineral micronutrients (Cu, Zn, Mn, Fe, Ni, Mo, B, Cl). Curr Opin Plant Biol 12:259-266

Hosoya LA, Sato Y, Fuller DQ (2010) Editorial: the archaeobotany of early rice agriculture in Asia. Archaeological and Anthropological Sciences 2:57-59

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961-967

Jolley VD, Brown JC, Blaylock MJ, Camp SD (1988) A role for potassium in the use of iron by plants. J Plant Nutr 11:1159-1175

Kang HM, Sul JH, Zaitlen NA, Kong S, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42:348-354

Kovach M, McCouch S (2008) Leveraging natural diversity: back through the bottleneck. Curr Opin Plant Biol 11:193-200

Kuper J, Llamas A, Hecht H, Mendel RR, Schwarz G (2004) Structure of the molybdopterin-bound Cnx1G domain links molybdenum and copper metabolism. Nature 430:803-806

Lahner B, Gong J, Mahmoudian M, Smith EL, Abid KB, Rogers EE, Guerinot ML, Harper JF, Ward JM, McIntyre L, Schroeder JI, Salt DE (2003) Genomic scale profiling of nutrient and trace elements in Arabidopsis thaliana. Nat Biotech 21:1215-1221

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez R (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-2948

Li G, Nunes L, Wang Y, Williams PN, Zheng M, Zhang Q, Zhu Y (2013) Profiling the ionome of rice and its use in discriminating geographical origins at the regional scale, China. Journal of Environmental Sciences 25:144-154

Liakat Ali M, McClung AM, Jia MH, Kimball JA, McCouch SR, Georgia CE (2011) A rice diversity panel evaluated for genetic and agro-morphological diversity between subpopulations and its geographic distribution. Crop Sci 51:2021-2035

Lowry DB, Sheng CC, Zhu Z, Juenger TE, Lahner B, Salt DE, Willis JH (2012) Mapping of ionomic traits in Mimulus guttatus reveals Mo and Cd QTLs that colocalize with *MOT1* homologues. PloS one 7:e30730

Ma J, Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. Proc Natl Acad Sci U S A 101:12404-12410

Maathuis FJ (2009) Physiological functions of mineral macronutrients. Curr Opin Plant Biol 12:250-258

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD (2007) The extent of linkage disequilibrium in rice (Oryza sativa L.). Genetics 177:2223-2232

Matsumoto TK, Pardo JM, Takeda S, Bressan RA, Hasegawa PM (2001) Tobacco and Arabidiopsis *SLT1* mediate salt tolerance of yeast. Plant Mol Biol 45:489-500

McCouch SR, McNally KL, Wang W, Hamilton RS (2012) Genomics of gene banks: A case study in rice. Am J Bot 99:407-423

McCouch SR, Zhao K, Wright M, Tung CW, Ebana K, Thomson M, Reynolds A, Wang D, DeClerck G, Ali ML, Others (2010) Development of genome-wide SNP assays for rice. Breed Sci 60:524-535

McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proceedings of the National Academy of Sciences 106:12273-12278

Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, Huang P, Jackson S, Schaal BA, Bustamante CD (2011) Molecular evidence for a single evolutionary origin of domesticated rice. Proceedings of the National Academy of Sciences 108:8351-8356

Morrissey J, Baxter IR, Lee J, Li L, Lahner B, Grotz N, Kaplan J, Salt DE, Guerinot ML (2009) The ferroportin metal efflux proteins function in iron and cobalt homeostasis in Arabidopsis. The Plant Cell Online 21:3326-3338

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu.Rev.Plant Biol. 59:651-681

Norton GJ, Dasgupta T, Islam MR, Islam S, Deacon CM, Zhao F, Stroud JL, McGrath SP, Feldmann J, Price AH (2010) Arsenic influence on genetic variation in grain trace-element nutrient content in Bengal Delta grown rice. Environ Sci Technol 44:8284-8288

Norton GJ, Deacon CM, Xiong L, Huang S, Meharg AA, Price AH (2010) Genetic mapping of the rice ionome in leaves and grain: identification of QTLs for 17 elements including arsenic, cadmium, iron and selenium. Plant Soil 329:139-153

Norton G, Duan G, Lei M, Zhu Y, Meharg A, Price A (2012) Identification of quantitative trait loci for rice grain element composition on an arsenic impacted soil: influence of flowering time on genetic loci. Ann Appl Biol 161:46-56

Ohnishi J, Flügge U, Heldt HW, Kanai R (1990) Involvement of Na in Active Uptake of Pyruvate in Mesophyll Chloroplasts of Some C4 Plants Na /Pyruvate Cotransport. Plant Physiol 94:950-959

Oka H (1988) Origin of cultivated rice. Japan Scientific Societies Press

Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD (2006) Selection under domestication: evidence for a sweep in the rice waxy genomic region. Genetics 173:975-983

Pearson K (1896) Mathematical Contributions to the Theory of Evolution.--On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. Proceedings of the Royal Society of London 60:489-498

Pilon-Smits EA, Quinn CF, Tapken W, Malagoli M, Schiavon M (2009) Physiological functions of beneficial elements. Curr Opin Plant Biol 12:267-274

Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y, Ito A, Ishii T, Miyashita NT, Terauchi R (2007) Large-scale DNA polymorphism study of Oryza sativa and O. rufipogon reveals the origin and divergence of Asian rice. Theor Appl Genet 114:731-743

Rampey RA, Woodward AW, Hobbs BN, Tierney MP, Lahner B, Salt DE, Bartel B (2006) An Arabidopsis basic helix-loop-helix leucine zipper protein modulates metal homeostasis and auxin conjugate responsiveness. Genetics 174:1841-1857

Ross-Ibarra J, Morrell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proceedings of the National Academy of Sciences 104:8641-8648

Rus A, Baxter I, Muthukumar B, Gustin J, Lahner B, Yakubova E, Salt DE (2006) Natural variants of AtHKT1 enhance Na accumulation in two wild populations of Arabidopsis. PLoS genetics 2:e210

Sagart L (2011) How many independent rice vocabularies in Asia?. Rice 4:121-133

Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, Wakimoto H, Yang C, Iwamoto M, Abe T (2013) Rice Annotation Project Database (RAP-DB): An integrative and interactive database for rice genomics. Plant and Cell Physiology 54:e6-e6

Salt DE, Baxter I, Lahner B (2008) Ionomics and the study of the plant ionome. Annual Review of Plant Biology 59:709-733

Sanchez DH, Piekenstain FL, Escaray F, Erban A, Kraemer U, Udvardi MK, Kopka J (2011) Comparative ionomics and metabolomics in extremophile and glycophytic Lotus species under salt stress challenge the metabolic pre‐adaptation hypothesis. Plant, Cell Environ 34:605-617

Sarkar N, Kim Y, Grover A (2009) Rice sHsp genes: genomic organization and expression profiling under stress and development. BMC Genomics 10:393

Sato Y, Namiki N, Takehisa H, Kamatsuki K, Minami H, Ikawa H, Ohyanagi H, Sugimoto K, Itoh J, Antonio BA (2013) RiceFREND: a platform for retrieving coexpressed gene networks in rice. Nucleic Acids Res 41:D1214-D1221

Schwarz G, Mendel RR (2006) Molybdenum cofactor biosynthesis and molybdenum enzymes. Annu.Rev.Plant Biol. 57:623-647

Shaff JE, Schultz BA, Craft EJ, Clark RT, Kochian LV (2010) GEOCHEM-EZ: a chemical speciation program with greater power and flexibility. Plant Soil 330:207-214

Shaibur MR, Kitajima N, Sugawara R, Kondo T, Huq SI, Kawai S (2006) Physiological and mineralogical properties of arsenic-induced chlorosis in rice seedlings grown hydroponically. Soil Sci Plant Nutr 52:691-700

Soetan K, Olaiya C, Oyewole O (2010) The importance of mineral elements for humans, domestic animals and plants: A review. African Journal of Food Science 4:200-222

Sweeney M, McCouch S (2007) The complex history of the domestication of rice. Annals of Botany 100:951-957

Tomatsu H, Takano J, Takahashi H, Watanabe-Takahashi A, Shibagaki N, Fujiwara T (2007) An Arabidopsis thaliana high-affinity molybdate transporter required for efficient uptake of molybdate from soil. Proceedings of the National Academy of Sciences 104:18807-18812

Tung C, Zhao K, Wright MH, Ali ML, Jung J, Kimball J, Tyagi W, Thomson MJ, McNally K, Leung H, Kim H, Ahn S, Reynolds A, Scheffler B, Eizenga G, McClung A, Bustamante C, McCouch SR (2010) Development of a Research Platform for Dissecting Phenotype–Genotype Associations in Rice (Oryza spp.). Rice. doi: 10.1007/s12284-010-9056-5

Vreugdenhil D, Aarts M, Koornneef M, Nelissen H, Ernst W (2004) Natural variation and QTL analysis for cationic mineral content in seeds of Arabidopsis thaliana. Plant, Cell Environ 27:828-839

Wang Y, Garvin DF, Kochian LV (2002) Rapid induction of regulatory and transporter genes in response to phosphorus, potassium, and iron deficiencies in tomato roots. Evidence for cross talk and root/rhizosphere-mediated signals. Plant Physiol 130:1361-1370

Watanabe T, Broadley MR, Jansen S, White PJ, Takada J, Satake K, Takamatsu T, Tuah SJ, Osaki M (2007) Evolutionary control of leaf element composition in plants. New Phytol 174:516-523

Williams L, Salt DE (2009) The plant ionome coming into focus. Curr Opin Plant Biol 12:247

Youens-Clark K, Buckler E, Casstevens T, Chen C, DeClerck G, Derwent P, Dharmawardhana P, Jaiswal P, Kersey P, Karthikeyan A (2011) Gramene database in 2010: updates and extensions. Nucleic Acids Res 39:D1085-D1094

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296:79-92

Zeng Y, Zhang H, Wang L, Pu X, Du J, Yang S, Liu J (2010) Genotypic variation in element concentrations in brown rice from Yunnan landraces in China. Environ Geochem Health 32:165-177

Zhao K, Tung C, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, Bustamante CD, McCouch SR (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nat Commun 2:467

Zhu J (2000) Genetic analysis of plant salt tolerance using Arabidopsis. Plant Physiol 124:941-948

CHAPTER 4:

OUTLINE OF FUTURE RESEARCH OBJECTIVES TO IDENTIFY THE GENETIC MECHANISMS UNDERLYING IONOMICS PHENOTYPES IN RICE AND TO UNDERSTAND THE RICE IONOME FROM A SYSTEMS BIOLOGY PERSPECTIVE

The long-term goal of this research is to provide a platform from which hypotheses can be generated about the physiological mechanisms that underlie the complex trait variation exhibited by the rice ionome. Investigating these mechanisms will require that several lines of evidence converge on a single locus in order to justify the expense of physiological experimentation. Some of these lines of evidence will simply provide validation of existing GWA hits. For instance, genetic mapping using bi-parental populations and the same phenotyping platform offers a powerful way of confirming GWA hits, or higher resolution haplotype analysis, coupled with in-depth evaluation of candidate genes underlying GWA peaks can help explain patterns observed in the GWA analysis, and sometimes even explain additional phenotypic variance. Other times, molecular investigation of genes already proposed to have an impact on the rice ionome need to be explored further (i.e., *OsSLT1* and *OsMOT1*). As these hits were not the sole determinants of the observed phenotypic variation, candidate genes associated with other, less significant hits should also be investigated to try to understand the full breadth of the genetic architecture that controls the ionome. Also, derivative phenotypes can be produced by principle component analysis or by

assembling other combinations of phenotypic covariates, such as root/shoot ratios, and these derived variables can be combined into a single metric that can be mapped, possibly identifying new QTLs, or sometimes co-localizing with existing regions of interest. In cases where multiple lines of evidence converge on a single locus, positional cloning may be justified to identify the key gene(s) involved. Alternatively, the diversity panel or segregating population(s) can be screened to identify recombinants of interest that help verify the effect of particular genomic regions on the phenotype. As genes and QTLs of interest are identified, markers can be developed to facilitate marker-assisted selection or other breeding objectives. Lastly, systems biology approaches are more and more important in the effort to fully understand the relationships that exist between constituents of the rice ionome. Understanding these relationships will add context to breeding activities related to mineral nutrition and generate additional hypothesis about which phenotypic covariates might share a common physiological mechanism.

Additional Mapping Efforts – Bi-parental

Validation of GWA peaks using bi-parental QTL analysis can help avoid Type 1 error. One hundred and fifty recombinant inbred lines developed from a cross between Azucena (*tropical japonica*) and IR64 (*indica*) (Ahmadi et al. 2005) have been phenotyped with the same ionomics phenotyping platform previously described. Once the phenotype data is fully adjusted using the Bayesian adjustment pipeline, QTL mapping will be undertaken with the expectation of identifying some QTL that co-localize with GWA peaks. The likelihood that this will occur in at least some instance

is high, given the similarity of growing conditions used in the two studies, and the fact that even when our GWA results were compared with QTLs identified by Norton et al. (2009), co-localization of GWA peaks and QTL was observed in several cases, despite the fact that Norton's study was conducted under field conditions. Using a subset of SNP markers identified on the IR64 x Azucena RIL population via Genotyping-By-Sequencing (GBS) (Elshire et. al. 2011, Spindel et al., 2013) and crudely adjusted phenotype data, six QTL for Co, Mg, Ni, and Pb have preliminarily been identified. Further refinement of the phenotype data and use of the full GBS dataset of 30K SNP markers will serve to validate these six QTL and ideally detect many more.

Additional Mapping efforts – GWA analysis

Part of the limitation of the GWA analysis described previously is that our sample size is low. A total of 394 lines were genotyped with the HDRA, and while this provides sufficient power to detect phenotype-genotype associations when the entire population is evaluated, subpopulation-specific evaluations typically lack power, due to a dramatic reduction in sample size. A new diversity panel (RDP2) is being developed at IRRI that will consist of an additional 1,500 varieties. Some of the subpopulation-specific panels will be comprised of 500 lines or more, and both RDP1 and RDP2 will be genotyped with the HDRA. The availability of this resource for the rice community will greatly enhance our ability to detect associations that are specific to one subpopulation or another. The primary challenge will be to collect high quality phenotypic data from so many diverse lines in a biologically meaningful and labor/cost effective way.

Additional Mapping efforts – Wilds

The *Oryza* genus harbors a tremendous amount of genetic variation (Wing et al. 2005), only a fraction of which is present in domesticated Asian rice (Kovach and McCouch 2008). The rice diversity panel described in this study actually consists of approximately 500 lines, of which ~400 belong to the cultivated *O. sativa* species, and an additional 100 accessions have been purified from wild *O. rufipogon* accessions from the IRRI gene bank. These wild accessions have also been genotyped using the HDRA (Jung, J and McCouch S, personal communication). This collection of wild ancestral materials will allow the rice community not only to survey additional allelic variation not present in *O. sativa*, but will have the power to resolve phenotype-genotype associations at a much finer level due to the rapid decay of linkage disequilibrium in *O. rufipogon* (Mather et al. 2007).

Evaluating haplotype variation using the HDRA and re-sequenced lines

There are many putative candidate genes identified in our ionomics study that are worthy of follow up. However, many of them will also be red herrings. Haplotype analysis and in-depth sequencing allows one to associate patterns of genomic variation across a candidate gene with patterns of phenotypic variation to validate or help reject a candidate gene hypothesis. For example, in addition to *OsSLT1* , the GWA analysis in the *aus* subpopulation for Na shoot content identified peaks that co-localized with five of the seven high affinity potassium transporters in the rice genome. There were also hits that failed to associate with a hypothesized candidate gene, and others that

proposed the involvement of as-yet uncharacterized genes in determining shoot Na uptake. As haplotype variation and candidate genes of interest are further examined, a picture will begin to emerge as to which haplotypes and genes contribute more or less to the phenotypic variance observed. As this information is discovered, graphical genotypes can be constructed based on haplotype variation across all candidate genes (and at peak SNPs for hits where no candidate exists) as the basis for developing hypotheses about the additivity of these loci. Where one locus fails to explain a major portion of the phenotypic variance, multiple loci acting together may accumulate enough of an effect to see a statistically significant association with the phenotype.

Cell Biology/Molecular physiology of *OsMOT1* and *OsSLT1*

One of the major discoveries to come out of our ionomics research is the identification of the candidate genes, *OsMOT1* and *OsSLT1* , as significantly associated with molybdenum concentrations and sodium concentrations, respectively, in rice shoot tissue. The molecular characterization of natural variation at these loci in rice and in other species is only just beginning (Baxter et al. 2008, Tomatsu et al 2007, Antoine et al. 2005, Matsumoto et al. 2001). Exploration of the molecular function of these genes in rice has the potential to contribute to the global understanding of the mechanisms underlying molybdenum and sodium partitioning within plants, generally, and the fortification and salt tolerance of rice, specifically. For instance, Baxter et al. (2008) determined that variation for gene expression of *AtMOT1* is the driving force governing molybdenum accumulation in the shoots of Arabidopsis. Our research demonstrated a significant association between highly divergent promoter sequences

and molybdate shoot concentrations in the *tropical japonica* subpopulation. In order to verify this hypothesis, experiments using quantitative RT PCR have begun to be carried out (Cornell, J and Maron L, personal communication) using some of the high and low molybdate accumulating lines from the RDP1.

Derivation and mapping of additional phenotypes

Our ionomics research demonstrated the power of principle components analysis to describe relationships between ionomic covariates. Because these relationships are strongly associated with specific subpopulations, we infer that there is genetic determinacy involved. Using the principle components themselves as phenotypes in a GWA analysis is one way to look for loci that act as master regulators of ionomics networks. Such peaks may be difficult to interpret biologically, but if they are significant, they would merit further investigation. It is possible that some hits discovered in this way may co-localize with hotspots for ionomic phenotypes in the genome, in which case additional testing would be warranted to try and understand why a seemingly innocuous region of the genome is so consistently associated with global patterns of ionomic distribution with in the plant.

Our study also identified several intriguing patterns of root/shoot distribution and proposed that the variation in these patterns is at some level associated with population structure. A full investigation into the genetic architecture of nutrient distribution would require the derivation of root/shoot ratio phenotypes and implementing those ratios directly as phenotypes in the GWA analysis. It is likely that some peaks would

co-localize, but important novel hits may also be identified that are specific to protein networks governing the complex cellular cascades that mobilize the movement of an element from the roots to other parts of the plant.

The use of targeted germplasm and admixed lines

Our ionomics analysis demonstrated the use of NILs as one tool to test hypotheses generated by GWA studies. Even where background introgressions are not completely purged, careful analysis and phenotyping can discover evidence that either supports or refutes a hypothesis. In this case, phenotyping lines for salt tolerance that carried well-defined introgressions across the region provided evidence of the potential function and breeding application of *OsSLT1* that GWA analysis alone could not provide. Recombinant inbred lines (RILs), MAGIC populations, and chromosome segment substitution lines (CSSLs) are all being developed in rice to empower the investigation of hypotheses about the genetics underlying quantitative variation and to explore new ways to utilize natural variation in plant genetics and breeding.

Mutlivariate analysis within and between phenotypes

With the development of diverse germplasm resources (RILs, CSSLs, MAGIC lines etc…) and continued expansion of the rice diversity panel, a critical core of publicly available, high quality shared rice germplasm is being made available to the research community. This is powerful because by leveraging a distributed phenotyping network, data can be collected on multitudes of different phenotypes. Just in the McCouch lab alone, root system architecture, panicle architecture, carbohydrate

remobilization, ionomics, cold tolerance, grain quality, agronomic and yield data are all being collected on the same set of lines. Network analyses such as principle components analysis, Gaussian graphical models, non-random skewers analysis, and gene co-expression analysis all have the power to integrate information within and across these complex phenotypes in order to make powerful predictions and inferences about biological systems and their relationship with environmental variables (explicit and hidden).

Onward and upward

The ionomics analysis described in this thesis is but a small manifestation of how new tools and technologies can be combined to generate and test novel hypotheses about genetic variation. As genetics and breeding become ever more information driven sciences, our ability to make progress will depend on our capacity to interpret the wealth of available phenotypic and genotypic information. A cautionary word is justified here, as there is potential to become overwhelmed by the sheer amount of data and for scientific progress to become mired in mountains of unorganized and otherwise useless data points. Just as biological systems require organization, hierarchy, and laws to orchestrate the complex biochemical networks that comprise an organism, our information systems must likewise be well structured, organized, and hierarchical to support innovation. Collecting enough data is no longer the challenge, rather the challenge is to be thoughtful about the data that gets collected, the way it is managed, databased, adjusted, and analyzed, and ultimately how it is accessed in order to maximize the benefit to scientific understanding and to society as a whole.

REFERENCES

Ahmadi, N., C. Dubreuil-Tranchant, B. Courtois, D. Foncéka, D. This, S. R. McCouch, M. Lorieux, J. C. Glaszmann, and A. Ghesquière. "New resources and integrated maps for IR64× Azucena, a reference population in rice." In *IRRI 5th International Rice Genetics Symposium and 3rd International Rice Functional Genomics Symposium, Manila, Philip*, pp. 19-23. 2005.

Antoine W, Stewart JM, De Los Reyes, Benildo G (2005) The rice homolog of the sodium/lithium tolerance gene functions as molecular chaperon in vitro. Physiol Plantarum 125:299-310

Baxter I, Muthukumar B, Park HC, Buchner P, Lahner B, Danku J, Zhao K, Lee J, Hawkesford MJ, Guerinot ML (2008) Variation in molybdenum content across broadly distributed populations of Arabidopsis thaliana is controlled by a mitochondrial molybdenum transporter (*MOT1*). PLoS Genetics 4:e1000004

Kovach M, McCouch S (2008) Leveraging natural diversity: back through the bottleneck. Curr Opin Plant Biol 11:193-200

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD (2007) The extent of linkage disequilibrium in rice (Oryza sativa L.). Genetics 177:2223-2232

Matsumoto TK, Pardo JM, Takeda S, Bressan RA, Hasegawa PM (2001) Tobacco and Arabidiopsis *SLT1* mediate salt tolerance of yeast. Plant Mol Biol 45:489-500

Spindel J, Wright M, Chen C, Cobb J, Gage J, Harrington S, Lorieux M, Ahmadi N, and McCouch S. Bridging the genotyping gap: using genotyping-by-sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theor Appl Genet (In Press).

Tomatsu H, Takano J, Takahashi H, Watanabe-Takahashi A, Shibagaki N, Fujiwara T (2007) An Arabidopsis thaliana high-affinity molybdate transporter required for efficient uptake of molybdate from soil. Proceedings of the National Academy of Sciences 104:18807-18812

Wing RA, Ammiraju JS, Luo M, Kim H, Yu Y, Kudrna D, Goicoechea JL, Wang W, Nelson W, Rao K (2005) The Oryza map alignment project: the golden path to unlocking the genetic potential of wild rice species. Plant Mol Biol 59:53-62